Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity

Karen O'Connell, Sue Jinks-Robertson,[1] and Thomas D. Petes[1]
Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, North Carolina 27710

ABSTRACT Two types of RNA:DNA associations can lead to genome instability: the formation of R-loops during transcription and the incorporation of ribonucleotide monophosphates (rNMPs) into DNA during replication. Both ribonuclease (RNase) H1 and RNase H2 degrade the RNA component of R-loops, whereas only RNase H2 can remove one or a few rNMPs from DNA. We performed high-resolution mapping of mitotic recombination events throughout the yeast genome in diploid strains of *Saccharomyces cerevisiae* lacking RNase H1 (*rnh1*Δ), RNase H2 (*rnh201*Δ), or both RNase H1 and RNase H2 (*rnh1*Δ *rnh201*Δ). We found little effect on recombination in the *rnh1*Δ strain, but elevated recombination in both the *rnh201*Δ and the double-mutant strains; levels of recombination in the double mutant were ~50% higher than in the *rnh201* single-mutant strain. An *rnh201*Δ mutant that additionally contained a mutation that reduces rNMP incorporation by DNA polymerase ε (*pol2-M644L*) had a level of instability similar to that observed in the presence of wild-type Pol ε. This result suggests that the elevated recombination observed in the absence of only RNase H2 is primarily a consequence of R-loops rather than misincorporated rNMPs.

KEYWORDS RNase H1; RNase H2; loss of heterozygosity; mitotic recombination; microarrays

RNA transcripts can stably associate with the DNA template during transcription, giving rise to R-loop structures containing an RNA:DNA hybrid and an unpaired DNA strand. Such R-loops can stall DNA replication forks and are a major source of recombination events that are stimulated by high levels of transcription (reviewed by Aguilera and Garcia-Muse 2012). Transcription-associated recombination (TAR) is generally strongest when the transcription machinery and a replication fork approach each other head on (Prado and Aguilera 2005), but there is also an orientation-independent component to such conflicts (Azvolinsky *et al.* 2009). Importantly, TAR is highly elevated when RNA processing is perturbed and R-loops accumulate, or when R-loop removal mechanisms are disabled. The RNA component of R-loops can be removed by the Sen1 RNA–DNA helicase (Mischo *et al.* 2011) or degraded by ribonuclease (RNase) H1 or RNase H2 (reviewed in Cerritelli and Crouch 2009), which are generally considered to be functionally redundant.

R-loops in yeast have been quantified genome-wide using immunoprecipitation with an antibody specific to DNA:RNA hybrids, followed by hybridization to a microarray or by DNA sequencing. This type of analysis has been done in wild type (Chan *et al.* 2014; El Hage *et al.* 2014), mRNA processing defective (Chan *et al.* 2014), RNase H-defective (Chan *et al.* 2014; El Hage *et al.* 2014), and RNA–DNA helicase (Sen1)-defective strains (Chan *et al.* 2014). In wild-type strains, Chan *et al.* (2014) observed RNA–DNA hybrid accumulation at Ty retrotransposons, near telomeres, within the ribosomal RNA (rRNA) gene cluster, and at highly transcribed GC-rich genes. El Hage *et al.* (2014) found hybrid accumulation in wild-type strains at highly transcribed genes, within the rRNA gene cluster, near tRNA genes, and near Ty retrotransposons. They also reported, however, that RNA:DNA accumulation in Ty retrotransposons was primarily a consequence of reverse transcription rather than R-loop formation. In strains lacking both RNases H1 and H2, both groups found increased R-loop accumulation, with the positions of the R-loops being similar to those observed in wild-type strains.

In addition to stable RNA:DNA association via R-loops, ribonucleotides (rNMPs) can be embedded into DNA during replication, either as remnants of Okazaki fragments or by

doi: 10.1534/genetics.115.182725
Manuscript received September 9, 2015; accepted for publication September 16, 2015; published Early Online September 22, 2015.
Supporting information is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.182725/-/DC1.
[1]Corresponding authors: Department of Molecular Genetics and Microbiology, 213 Research Dr., CARL Bldg., Duke University Medical Center, Durham, NC 27710. E-mail: tom.petes@duke.edu; and sue.robertson@duke.edu

direct incorporation (reviewed by Williams and Kunkel 2014). Despite the efficiency with which replicative DNA polymerases discriminate between rNTP and dNTP precursors, there are ~15,000 rNMPs inserted into yeast DNA during each round of replication, which translates into ~1 rNMP per 1000 nucleotides (Nick McElhinny *et al.* 2010b). rNMPs in budding yeast genomic DNA have recently been mapped to single-nucleotide resolution by several groups (Clausen *et al.* 2015; Koh *et al.* 2015; Reijns *et al.* 2015). In contrast to R-loops, which are removed by RNase H1 (encoded by *RNH1*) or RNase H2 (catalytic subunit encoded by *RNH201*), only RNase H2 is capable of removing of single rNMPs embedded in DNA (Cerritelli and Crouch 2009). In the absence of error-free removal by RNase H2, error-prone removal of single rNMPs can be initiated by Topoisomerase 1 (Top1), resulting in short deletions (Kim *et al.* 2011; Williams *et al.* 2013). Although genetic instability in the absence of RNase H1 and H2 is generally assumed to reflect persistent R-loops, pathways that promote DNA-damage bypass during replication are essential in the complete absence of RNase H activity (Lazzaro *et al.* 2012). The requirement for bypass activity suggests that contiguous tracts of rNMPs in genomic DNA may also contribute to genetic instability.

In addition to the specific case of recombination associated with highly elevated transcription, yeast strains with reduced RNase H activity generally exhibit elevated genomic instability. Haploid strains lacking RNase H2, for example, have increased recombination between direct repeats (Ii *et al.* 2011; Potenski *et al.* 2014). Top1 is required for this increase in recombination, suggesting that the initiating lesion is a Top1-mediated nick at an rNMP (Potenski *et al.* 2014). Contributions of RNase H activity to genome integrity have additionally been assessed by measuring artificial chromosome stability or loss of heterozygosity (LOH) on an endogenous yeast chromosome (Wahba *et al.* 2011). Mutations eliminating either RNase H1 or RNase H2 increased the rate of loss of an artificial chromosome, and lack of both enzymes had a synergistic effect. With regard to LOH, however, an increase was observed only in an *rnh1*Δ *rnh201*Δ double-mutant background.

It is clear that yeast strains lacking both RNase H1 and RNase H2 activity have elevated levels of recombination, but the relative contributions of R-loops and rNMPs have not been explored. In addition, most previous studies have used assays that are limited either to a single locus or a single chromosome arm without subsequent mapping of repair events. In the current study, we examine the location and distribution of LOH events that occur throughout the yeast genome or on the 1-Mb right arm of chromosome IV. Analyses were done in single-mutant *rnh1*Δ and *rnh201*Δ diploids, as well as in double-mutant *rnh1*Δ *rnh201*Δ diploids. We found that the *rnh1*Δ strain had levels of recombination indistinguishable from those in wild type, while LOH was elevated in an *rnh201*Δ strain. Experiments done under conditions of reduced rNMP incorporation into genomic DNA indicated that the elevated recombination in the *rnh201*Δ single-mutant background reflected primarily the accumulation of R-loops. In all LOH assays, instability was further increased in the *rnh1*Δ *rnh201*Δ double-mutant relative to the *rnh201*Δ single-mutant strain. These findings have relevance to human disease, where hypomorphic mutations in RNase H2 lead to the neurodegenerative disorder Aicardi-Goutiéres syndrome (Crow *et al.* 2006) and have been associated with systemic lupus erythematosus (Günther *et al.* 2015).

Materials and Methods

Strain construction

Haploid strains with various gene deletions were constructed by one-step transplacement in isogenic derivatives of W303-1A (genotype of derivative: *MAT***a** *ade2-1 can1-100 ura3-1 ho*::*hisG his3-11,15 leu2-3,112 trp1-1*) and YJM789 (genotype of derivative: *MAT*α *ade2-1 ho*::*hisG gal2 ura3*) genetic backgrounds using polymerase chain reaction (PCR)-generated DNA fragments containing a selectable drug-resistance marker. Deletions were confirmed using PCR. To introduce the *pol2-M644L* mutant allele, two-step transplacement was done using the *Age*I-digested plasmid p173-pol2-M644L (Nick McElhinny *et al.* 2010a).

LOH on chromosome IV was examined using hybrid diploid strains formed by mating W303-1A derivatives with YJM789 derivatives. Diploids had a *URA3* gene from *Kluyveromyces lactis* (*URA3-Kl*) inserted near the right telomere of chromosome IV (*Saccharomyces* Genome Database, SGD coordinate 14954320) on the W303-1A-derived homolog, and an *ADE2* gene inserted at the allelic position on the YJM789-derived homolog. For all experiments, at least two isolates of each diploid genotype were analyzed. In addition, for each diploid, at least 10 tetrads were dissected to confirm the correct genotypes.

A complete list of strains and the details of their construction are provided in Supporting Information, Table S1 and File S1. Plasmids and primers used in the strain construction are listed in Table S2 and Table S3, respectively. Table S4 summarizes which strains were used in different assays of LOH.

Accumulation of recombination events in subcultured cells

Mutant and wild-type strains were serially passaged on rich YPD medium (yeast, peptone, dextrose: 1% yeast extract, 2% bacto-peptone, 2% dextrose; 2% agar for plates) to allow accumulation of recombination events genome-wide. In each passage, a strain was grown from a single cell into a colony. This procedure was repeated 10 or 20 times, corresponding to ~250 and ~500 cell divisions, respectively. Genomic DNA from subcultured strains was examined using whole-genome single-nucleotide polymorphism (SNP) arrays.

Rates of LOH on chromosome IV

Single colonies grown on rich medium were diluted and plated onto solid synthetic dextrose medium supplemented

with all amino acids and bases except arginine. The concentration of adenine was reduced to 10 μg/ml to allow the Ade$^-$ sectors to develop their red color. Each plate contained ~1000 colonies. The plates were screened for red/white sectored colonies using a dissecting microscope. Single colonies derived from the red and white sides of each sectored colony were isolated for subsequent phenotypic and physical analysis. To form a red/white sectored colony, the crossover must occur in the first division after the cells are plated. Thus, the frequency of red/white colonies divided by the total number of colonies is also the rate of sectored colony formation per cell division.

To produce a reciprocal crossover (RCO) using the sectoring assay, both daughter cells containing the recombination products must be viable. For each strain, we examined the daughter-cell viability by following the division of 88 unbudded cells on solid medium. After ~60 min of growth at 30°, the mother and daughter cells were separated through micromanipulation. After 2 days at 30°, the growth of individual cells into colonies was assessed. For all strains, at least 75% of the unbudded cells produced two viable daughter cells.

In separate experiments, we measured the rate of loss of the *URA3-Kl* marker by monitoring the rate of colonies resistant to 5-fluoro-orotate (5-FOA) as described in detail in File S1. To measure rates, we determine the frequencies of 5-FOA^R derivatives in 10–20 individual cultures. These frequencies are converted to rates of 5-FOA resistance (the number of 5-FOA^R isolates generated per cell division) using the method of the median (Lea and Coulson 1949). At least two independently derived diploid isolates were used for each estimate. Confidence intervals were obtained as described in Altman (1991).

Microarray hybridization and analysis

Diploids used to monitor LOH were heterozygous for ~55,000 SNPs. Chromosome IV-specific microarrays monitored ~1000 SNPs on the right arm of chromosome IV, and whole-genome microarrays monitored ~13,000 SNPs distributed throughout the genome (St Charles *et al.* 2012, 2013). On our custom Agilent microarrays, each SNP was represented by at least four 25-base oligonucleotides: probes identical to the Watson and Crick strands of the W303-1A-derived SNP and probes identical to the Watson and Crick strands of the YJM789-derived SNP. The sequences used for the oligonucleotides for these arrays are described in St Charles *et al.* (2012, 2013).

DNA samples derived from the red or white sides of sectored colonies were labeled with a Cy5-tagged fluorescent nucleotide (St Charles *et al.* 2012). DNA from a wild-type heterozygous control sample was labeled with a Cyanine3 (Cy3)-tagged fluorescent nucleotide. The labeled control and experimental DNAs were then mixed and hybridized to the microarrays. The hybridization levels for each Cy3- and Cy5-labeled sample were determined using a GenePix scanner. The ratio of Cy5 to Cy3 fluorescence (*i.e.*, the ratio of medians) was normalized for each array (St Charles *et al.* 2012). If the normalized ratio of medians of a given sample for both W303-1A- and YJM789- derived oligonucleotides was 1, the locus is heterozygous. LOH events result in a normalized ratio of medians >1.5 for oligonucleotides representing one homolog and a signal <0.5 for oligonucleotides representing the other homolog. We also required that at least two adjacent SNPs reflect LOH to be included in the dataset. In addition to detecting LOH events, the SNP microarrays revealed deletion and duplication events (reduced or elevated hybridization levels for SNPs from one homolog with no alteration in the level of hybridization of SNPs on the other homolog) and ploidy changes (trisomy).

For each SNP on the microarray, our analysis allows us to conclude whether the analyzed strain is homozygous for the W303-1A-derived SNP, the YJM789-derived SNP, or retains heterozygosity. The SGD coordinates for LOH events, deletions/duplications, and ploidy changes in the subcultured strains are shown in Table S5, Table S6, and Table S7, respectively. SGD coordinates for LOH events in sectored colonies are in Table S8. The patterns of LOH, deletions/duplications, and aneuploidy for subcultured strains are shown schematically in Figure S1, Figure S2, and Figure S3; the location of LOH events on the chromosomes in subcultured strains is summarized in Figure S4, Figure S5, and Figure S6. The patterns of LOH for sectored colonies are shown schematically in Figure S7.

Association of genome features with recombination events

The transition between heterozygous and homozygous SNPs should contain the site of the recombination-initiating lesion (St Charles and Petes 2013; Yin and Petes 2013). Breakpoint regions in subcultured mutant strains were examined to determine if specific chromosome elements (for example, autonomously replicating sequence, ARS elements) were overrepresented at the breakpoints. For interstitial LOH events (gene conversions), we used association windows that extended between the two heterozygous flanking SNPs that were closest to the homozygous SNPs at the termini of the conversion tracts. The association windows for terminal LOH events were defined by including the sequences located 10 kb centromere-proximal and 10 kb centromere-distal to the first homozygous SNP. For RCO events associated with sectored colonies, we used association windows that extended from the centromere-proximal heterozygous SNP of the event to the most centromere-distal homozygous SNP of the event. Other details of the association analysis are described in File S1.

The references for the locations of the chromosome elements in the genome are in Table S9, and the number of elements per genomic microarray are in Table S10. After defining the association windows for LOH events (as described above), we tallied the number of specific genomic elements within the association windows for each genotype for each type of experiment (whole-genome or chromosome IV-specific analyses). The total numbers of genomic elements within and outside of the association windows were compared

to the expected numbers by χ^2 analysis. The expected numbers were based on the total numbers of elements in the genome, and the amount of DNA inside and outside of the association windows for all strains analyzed (Song *et al.* 2014). Because multiple genome features were tested for association with recombination windows, we applied a correction of the *P*-value (Hochberg and Benjamini 1990). The association analyses for the subcultured strains and the sectored colonies are in Table S11 and Table S12, respectively.

Calculations of gene conversion tract lengths

Gene conversion tract lengths were calculated for RCO events in the sectoring assay and for interstitial LOH events in subcultured strains. Tract lengths were measured by calculating the distances between the midpoints of heterozygous to homozygous transitions on the farthest edges of both sides of the tracts.

Data availability

Strains are available upon request. The contents of File S1, Table S1, Table S2, Table S3, Table S4, Table S5, Table S6, Table S7, Table S8, Table S9, Table S10, Table S11, and Table S12, and Figure S1, Figure S2, Figure S3, Figure S4, Figure S5, Figure S6, and Figure S7 are described in the text. Microarray data are available at GEO with the accession number GSE73334.

Results

To assess the contributions of incorporated ribonucleotides (rNMPs) and R-loops to genome instability and recombination in RNase H-defective strains, we compared genome instability in the following six diploid strains: wild type (KO198), *rnh1*Δ (KO73/KO187), *rnh201*Δ (KO75/KO188/KO135), *pol2-M644L* (KO234), *rnh201*Δ *pol2-M644L* (KO244), and *rnh1*Δ *rnh201*Δ (KO5/KO132/KO189); the complete genotypes of all strains are given in Table S1. The *rnh1*Δ and *rnh201*Δ mutants lack RNase H1 and the catalytic subunit of RNase H2, respectively (reviewed in Cerritelli and Crouch 2009). The *pol2-M644L* allele encodes a mutant form of the catalytic subunit of DNA polymerase ε (Pol ε) that results in the insertion of 70% fewer rNMPs than the wild-type enzyme (Nick McElhinny *et al.* 2010a). Genome instability in mutant and wild-type diploid strains was monitored using two approaches. The first was to examine LOH throughout the genome in cells that were serially passaged for ~500 divisions (20 cycles of growth from a single cell to a colony). The second approach involved identification of cells that had undergone a crossover on the right arm of chromosome IV. The locations of LOH events were determined by microarray analysis. Details of these systems are described further below.

Genome-wide mapping of LOH events in subcultured RNase H-defective strains

Experiments utilized diploid strains that were derived by mating two sequence-diverged haploids (W303-1A and YJM789) and were heterozygous for ~55,000 single-nucleotide SNPs (St Charles *et al.* 2012). We developed oligonucleotide-based microarrays that detect LOH for ~13,000 SNPs distributed throughout the genome, allowing mapping of events to ~1-kb resolution. Each heterozygous SNP on the microarray was represented by at least four 25-base oligonucleotides: two identical to the W303-1A form of the SNP and two identical to the YJM789 form. DNA from the serially passaged strains was labeled with Cy5-fluorescent dNTPs and mixed with DNA from a wild-type heterozygous control strain, which was labeled with Cy3-fluorescent dNTPs. The combined sample was then hybridized to a whole-genome microarray and the hybridization signal of DNA from the passaged strain was normalized to that from the heterozygous control strain (details in *Materials and Methods*). A normalized ratio of about one for both the YJM789 and W303-1A SNP indicated that the subcultured sample had maintained heterozygosity at that position. LOH, as a consequence of a crossover, resulted in an elevation in the hybridization signal for strain-specific SNPs derived from one homolog and a reduction in the signal of SNPs derived from the other homolog. Heterozygous deletions or duplications resulted in a loss or an increase, respectively, in the hybridization of SNPs specific for one strain with no alteration in the hybridization level of the other strain-specific SNPs. Changes in ploidy were also readily detected.

The four common classes of events observed by this analysis are shown in Figure 1. In this depiction, red and blue lines represent levels of hybridization to W303-1A- and YJM789-derived SNPs, respectively. An interstitial LOH event is shown in Figure 1A; from previous studies (St Charles *et al.* 2012; St Charles and Petes 2013; Yin and Petes 2013), we know that these events reflect gene conversion, the nonreciprocal transfer of sequences from one homolog to the other. In this example, sequences of the YJM789 copy of chromosome II were copied into the W303-1A copy of chromosome II at two different locations. Figure 1B shows a duplication of sequences derived from the W303-1A copy of chromosome XIII; such events are often generated by unequal recombination between two nonallelic Ty elements or other repeated genes (Song *et al.* 2014). Figure 1C shows a terminal LOH event that includes most of the left arm of chromosome XI. This pattern of LOH can result from a crossover or break-induced replication (BIR) event between homologs. Lastly, Figure 1D shows the pattern of hybridization expected for chromosome V trisomy with a duplication of the YJM789-derived homolog. No chromosome loss events were observed in our experiments. Schematic depictions of all of the LOH patterns observed in these strains are shown in Figure S1 (terminal and interstitial LOH), Figure S2 (deletions/duplications), and Figure S3 (aneuploidy).

We determined the numbers of events of various classes described above after 10–21 colonies of each genotype were subcultured 20 times, which corresponds to ~500 generations. These data are summarized in Figure 2. The average number of events (sum of all four classes) was about one per

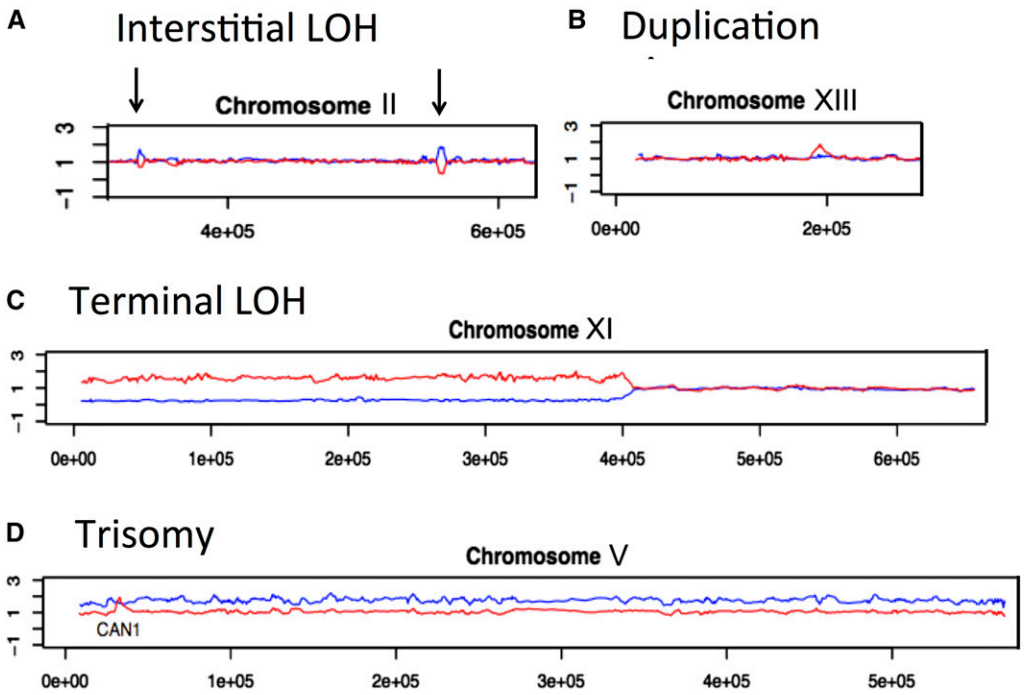

Figure 1 Classes of events found in subcultured strains. In each panel, the *x*-axis represents the chromosome coordinate, and the *y*-axis represents the ratio of hybridization medians. The red and blue lines reflect the hybridization signal of the W303-1A-specific- and YJM789-specific oligonucleotides, respectively. The events in this figure are from the genome of a single passaged isolate of the *rnh1*Δ *rnh201*Δ double mutant.

strain in the wild-type, *rnh1*Δ, and *pol2-M644L* strains. There was no difference in the number of events for these three genotypes (Wilcoxon rank sum test, $P > 0.05$). In contrast, the average number of events was elevated three- to fourfold in the *rnh201*Δ single- and *rnh201*Δ *pol2-M644L* double-mutant strains relative to wild type, and about sevenfold in the *rnh1*Δ *rnh201*Δ double mutant. Differences in the average number of total events per strain primarily reflected differences in the number of terminal and interstitial LOH events. By the Wilcoxon rank sum test, all three of these strains had significantly ($P < 0.001$) more genome alterations than in wild type, and the *rnh1*Δ *rnh201*Δ strain had significantly ($P < 0.001$) more rearrangements than either the *rnh201*Δ or the *rnh201*Δ *pol2-M644L* strain.

The lack of an increase in recombination in the *rnh1*Δ strain, the greater instability in the *rnh201*Δ mutant, and the even greater instability in the *rnh1*Δ *rnh201*Δ double mutant could be explained in two ways. First, it is possible that the recombinogenic lesions primarily reflect persistent R-loops and RNase H2 is the primary enzyme involved in their removal, with RNase H1 playing a back-up role. The alternative possibility is that the primary recombinogenic lesions are misincorporated rNMPs, with RNase H2 acting as the most important enzyme in their removal and RNase H1 having a relatively minor effect. Data obtained with the *rnh201*Δ *pol2-M644L* double-mutant strain argue against the second possibility. As described previously, the *pol2-M644L* allele encodes a mutant Pol ε that incorporates only one-third as many rNMPs as the wild-type enzyme (Nick McElhinny *et al.* 2010a). If rNMPs are the major recombinogenic lesion in the *rnh201*Δ single mutant, then the level of instability in the *rnh201*Δ *pol2-M644L* double mutant should decrease to one-third of the level observed in *rnh201*Δ single mutant. No significant decrease in instability was observed ($P = 0.37$), however, indicating that rNMPs embedded into DNA by Pol ε are not the major source of instability in the *rnh201*Δ single-mutant strain. The most likely interpretation of the data in Figure 2 is that the hyperrecombination (hyperrec) phenotype is driven primarily by persistent R-loops, which are present at inconsequential levels in the *rnh1*Δ single mutant, high levels in the *rnh201*Δ single mutant, and even higher levels in the *rnh1*Δ *rnh201*Δ mutant.

In addition to determining the number of LOH events and other genomic alterations, we used the microarray data to map the location of the transitions between heterozygous and homozygous SNPs. The coordinates of the SNPs adjacent to the transitions of LOH events are presented in Table S5, Table S6, and Table S7. These coordinates are based on the June 2008 (SGD/sacCer2) version of the SGD available on the University of California Santa Cruz Genome Browser (http://genome.ucsc.edu/).

Several types of analyses were performed with these data. First, simple interstitial LOH events (class b1-2 from Figure S1) were used to estimate the length of the gene conversion tract associated with each event. The conversion events that produce interstitial LOH are both conversion events that are unassociated with crossovers and conversion events in which the recombinant chromosomes cosegregate into the same cell. Tract size was estimated by averaging the distance between the closest heterozygous sites flanking the transition (the maximum tract size) and the sites of homozygous SNPs closest to the transitions (the minimum tract size). A summary of conversion tract sizes is in Table S5. The median conversion tract lengths (in parentheses) for each strain were: *rnh201*Δ (11.8 kb), *rnh1*Δ *rnh201*Δ (11.8 kb), and *rnh201*Δ *pol2-M644L* (14.2 kb). The median length of conversion

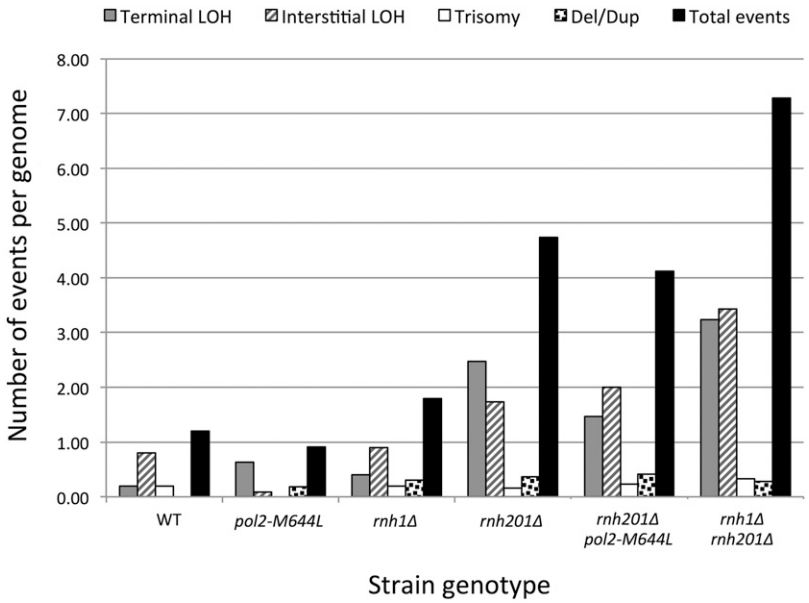

Figure 2 Genome instability of subcultured wild type, *pol2-M644L*, *rnh1*Δ, *rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ strains. The total number of events per genome is indicated for each genotype as a black bar. For wild-type, *pol2-M644L*, *rnh1*Δ, *rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ strains, 10, 12, 10, 19, 17, and 22 genomes were analyzed by whole-genome microarrays, respectively. Since the wild-type strain was subcultured 10 times, and the mutant strains were subcultured 20 times (~500 cell divisions), we doubled the number of events observed in the wild-type strain to make the comparisons among the strains valid.

events in the wild-type strain (based on only four events) was 12.7 kb, similar to the lengths observed in the mutant strains.

In the second type of analysis, we mapped the location of interstitial and terminal LOH events in yeast genome to search for potential hotspots. As shown in Figure S4, Figure S5, and Figure S6, LOH events in the *rnh201*Δ, *rnh1*Δ *rnh201*Δ, and *rnh201*Δ *pol2-M644L* mutants were widely distributed throughout the genome with little evidence for strong hotspots of recombination. For strains with large numbers of LOH events (*rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ), we also examined regions near the recombination breakpoints for enrichment of various chromosome elements. The rationale for this analysis is that the breakpoints are likely to be located at or near the site of the recombination-initiating DNA lesion. For each simple terminal and interstitial LOH event, we calculated an association window likely to contain the relevant lesion. For interstitial events, the association windows were the DNA sequences between the heterozygous SNPs that most closely flanked the LOH region (details are provided in *Materials and Methods* and File S1). For terminal LOH events, we used a 20-kb window, extending 10 kb centromere-proximal and centromere-distal from the homozygous SNP located at the transition point. After calculating the association window for each event, we determined the total number of each chromosome element/motif located within these windows. Based on the total number of an element in the genome and the fraction of genome located within and outside of the windows, we calculated an expected number of elements within and outside of the windows. We then compared the observed numbers of events to the expected numbers by χ^2 analysis, correcting the $\underline{P}$-values for multiple comparisons (Hochberg and Benjamini 1990). The list of the elements used in our analysis and the expected numbers of such elements are given in Table S9 and Table S10, respectively. The results of this analysis for the subcultured strains are in Table S11.

In our previous studies, breakpoints of spontaneous recombination events in wild-type strains were enriched for replication fork-stalling motifs (St Charles and Petes 2013); a similar enrichment was observed for events associated with low levels of DNA Pol α (Song *et al.* 2014). In the current analysis, we first looked for overrepresented or underrepresented genomic elements in the subcultured *rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ strains. Based on results of several other groups, we expected overrepresentation of regions that accumulate R-loops and/or rNMPs among the recombination breakpoints. R-loop formation is promoted by high GC content, the ability to form G4 quadruplex structures on the nontranscribed strand, high levels of transcription, and unusually long genes (reviewed by Aguilera and Garcia-Muse 2012; Hamperl and Cimprich 2014). We found no significant enrichment for these factors (Table S11). Additionally, genomic sites of R-loop accumulation have been measured directly in wild-type and *rnh1*Δ *rnh201*Δ strains (Chan *et al.* 2014; El Hage *et al.* 2014). Surprisingly, these sites also were not enriched at the LOH breakpoints in our study. An exception may be the ribosomal RNA gene tandem array, where two independent *rnh1*Δ *rnh201*Δ diploids had LOH events before subculturing began, indicating a high level of instability. We also found that one of 19 *rnh201*Δ and three of 17 *rnh201*Δ *pol2-M644L* subcultured colonies had LOH events in the rDNA; there were none among the 10 wild-type colonies analyzed. This difference, however, is not significant by the Fisher exact test ($P = 0.56$).

We also looked for overrepresentation of sites of RNA polymerase II subunit Rbp3 accumulation during S phase (Fachinetti *et al.* 2010) and binding sites for the Rrm3 helicase (Azvolinsky *et al.* 2009) at LOH breakpoints. No significant associations were observed for any of three mutant

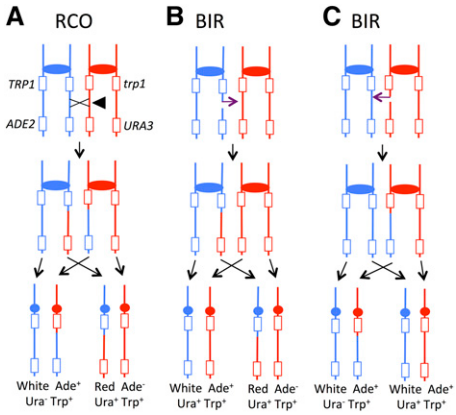

Figure 3 Mechanisms of repair leading to terminal LOH events in the red/white sectoring system and in 5-FOA-resistant colonies. Blue and red lines represent the YJM789-derived and W303-derived chromosomes, respectively. Rectangles and circles represent genes and centromeres, respectively. In two of the three types of repair in this figure, a red/white sectored colony is formed. (A) A double-strand break (DSB) leads to a reciprocal crossover (RCO) event. After the first mitotic division following the crossover, the sectored colony has reciprocal LOH products in the white, Ura$^-$ and the red Ura$^+$ sides of the sector. (B) A DSB on the YJM789-derived homolog repairs through break-induced replication (BIR) of the W303-1A-derived homolog, leading to a nonreciprocal terminal LOH event. Both sides of the red/white sector are Ura$^+$. (C) A DSB on the W303-1A-derived homolog repairs by BIR, using the YJM789-derived homolog as a template. Although no red/white sectored colony is formed, one half of the colony is Ura$^-$ and the other half is Ura$^+$.

strains (Table S11). Lastly, we calculated whether LOH events located within 25 kb of the telomere were overrepresented in the datasets. We found an overrepresentation of LOH events near the telomere in the *rnh201*Δ single mutant (χ^2 analysis; *P*-value of 0.002), but not in the *rnh201*Δ *pol2-M644L* or *rnh1*Δ *rnh201*Δ double mutant.

Frequency of recombination events on the right arm of chromosome IV

In addition to examining patterns of unselected LOH in subcultured mutant strains, we also employed a system that identifies cells that have undergone a RCO or BIR event on the ~1 Mb right arm of chromosome IV. As will be discussed below, this system allows inference of whether the recombination-initiating lesion occurred in G_1 or S/G_2 of the cell cycle. In the same hybrid genetic background used for the whole-genome analyses, we constructed diploids in which YJM789-derived copy of chromosome IV had an insertion of *ADE2* near the telomere of the right arm. The W303-1A-derived homolog had an insertion of *URA3* at the same position (Figure 3). The diploid was also homozygous for the *ade2-1* allele at the *ADE2* locus on chromosome XV. Strains without a wild-type *ADE2* gene on chromosome IV form red colonies.

An RCO event between the *ADE2/URA3* insertions and *CEN4* can produce a red/white sectored colony in which the red side of the sector is Ura$^+$ and the white side is Ura$^-$ (Figure 3A). In contrast, a BIR event initiated by a break on the YJM789-derived homolog produces a red/white sectored colony in which both sides are Ura$^+$ (Figure 3B). It should be noted that a red/white sectored colony in which both sectors are Ura$^+$ also can result from loss of the YJM789-derived chromosome, producing the red side of the sector. In this case, however, the red side of the sector would be Trp$^-$. Of 173 red/white Ura$^+$/Ura$^+$ sectors examined, the red sector was Trp$^+$ in all but one, indicating that chromosome loss does not significantly contribute to sector formation. It should be noted that if a BIR-initiating break occurs on the W303-1A-derived homolog, a red/white sectored colony is not formed (Figure 3C).

Using this system, we quantitated RCO and BIR events on the right arm of chromosome IV in two different ways. First, we counted red/white sectored colonies relative to total colonies in the following diploid strains: wild type, *rnh1*Δ, *pol2-M644L*, *rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ. Purified colonies from the red and white sectors were checked to determine whether they were Ura$^+$ or Ura$^-$ to distinguish between BIR and RCO events. A summary of this analysis is shown in Figure 4. When normalized to the rate of sectored colonies observed in wild type, the rate of sectored colonies was 1.0 in *rnh1*Δ, 1.6 in *pol2-M644L*, 3.6 in *rnh201*Δ, 4.0 in *rnh201*Δ *pol2-M644L*, and 5.7 in *rnh1*Δ *rnh201*Δ. The sectoring results were in broad agreement with genome-wide analysis done using subcultured colonies: *rnh1*Δ and *pol2-M644L* were indistinguishable from wild type ($P = 0.92$ and $P = 0.16$, respectively; contingency χ^2), sectoring was elevated in *rnh201*Δ relative to wild type ($P < 0.001$), the presence of the *pol2-M644L* allele did not reduce sectoring in the *rnh201*Δ background ($P = 0.60$), and sectoring was elevated in the *rnh1*Δ *rnh201*Δ double mutant relative to the *rnh201*Δ single mutant ($P = 0.002$). Relative to wild type, there were changes in the distribution between BIR and RCO events in the *rnh201*Δ *pol2-M644L* and *rnh1*Δ *rnh201*Δ strains, where the BIR events were a larger fraction of LOH events. This difference was statistically significant only for the *rnh201*Δ *pol2-M644L* strain ($P = 0.04$ by Fisher exact test). In our analysis of subcultured colonies, it was not possible to distinguish between RCO and BIR events.

Our estimates of the rates of sectored colonies were based on a nonselective screening procedure and involved a relatively small number of sectored colonies, from 14 for the wild type to 126 for the *rnh201*Δ strain. Since both BIR events that initiate on the W303-1A homolog and RCO events can result in derivatives that are Ura$^-$ and, therefore, selectable on medium containing 5-FOA (Figure 3, A and C), we also determined the rates of 5-FOA^R colonies in the same strains used for the red/white sector analysis. These results (Figure 5) are also in good agreement with the LOH data from whole-genome analysis of the subcultured mutants (Figure 2). The wild-type, *rnh1*Δ, and *pol2-M644L* strains had similar rates of LOH on the right arm of chromosome IV (~5×10^{-5}/cell division). The *rnh201*Δ and *rnh201*Δ *pol2-M644L* strains had rates that were elevated ~5-fold above wild type, and the *rnh1*Δ *rnh201*Δ double mutant had a rate ~10-fold higher

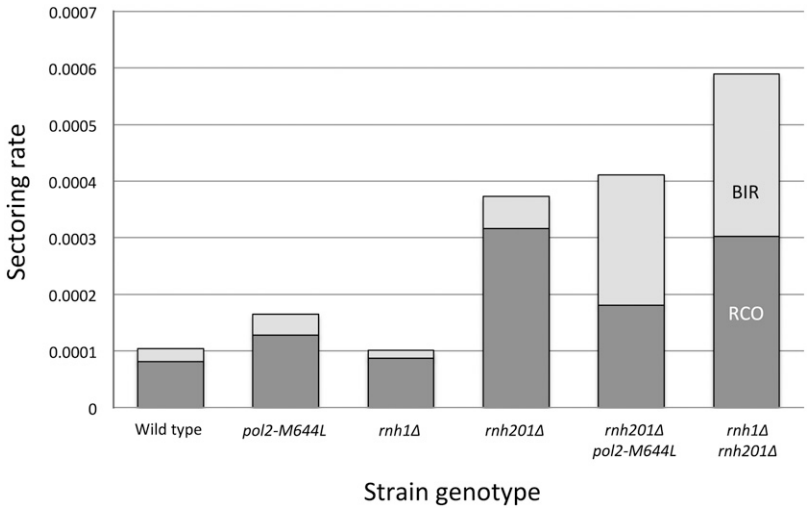

Figure 4 Rate of red/white sector formation in RNase H-defective strains. The numbers of sectors among total colonies screened for wild type, *pol2-M644L*, *rnh1*Δ, *rnh201*Δ, *rnh201*Δ *pol2-M644L*, and *rnh1*Δ *rnh201*Δ strains were 14/134864, 44/266267, 22/218008, 126/337864, 59/143669, and 78/132302, respectively. Dark and light gray bars correspond to RCO and BIR events, respectively, among sectored colonies.

than wild type. Though the rates of LOH were similar for the *rnh201*Δ and *rnh201*Δ *pol2-M644L* strains, it should be noted that there was a significant, ~20% decrease in the double mutant ($P = 0.016$ by Mann–Whitney test).

Mapping RCO events on the right arm of chromosome IV

The red and white sides of sectored colonies were analyzed using chromosome IV-specific microarrays, which allows the position of crossovers and their associated gene conversion tracts to be mapped (St. Charles *et al.* 2012; St. Charles and Petes 2013). For RCO events unassociated with gene conversion, the transition between heterozygous and homozygous SNPs occurs at exactly the same position in both sectors, indicated by the dotted line in Figure 6A. Gene conversion events in which one sister chromatid is broken and subsequently repaired using a nonsister chromatid result in an LOH pattern in which the breakpoints between heterozygous and homozygous SNPs are at different positions in the two sectors (Figure 6B). In the region boxed in Figure 6B, three of the chromatids have SNPs specific to W303-1A and one has SNPs specific to YJM789. This type of conversion is, therefore, defined as a 3:1 conversion. We infer that the DNA lesion that initiated the crossover likely occurred in S or G_2 of the cell cycle because only one sister chromatid received information from the homologous donor chromosome. Another common LOH pattern observed for spontaneous crossovers on chromosome IV is a 4:0 conversion (Figure 6C). This class of event reflects the repair of two sister chromatids broken at approximately the same position. The 4:0 pattern of conversion likely reflects double-strand break (DSB) formation in G_1 that is subsequently replicated to give two broken chromatids (Lee and Petes 2010). Microarray analyses of the red and white sectors with these patterns are in Figure 7. In the boxed region in Figure 7B, in the red sector, the hybridization signal is about one, indicating one copy of W303-1A-derived SNPs and one copy of YJM789-derived SNPs; in the white sector, the ratio of hybridization indicates that there are two copies of YJM789-derived SNPs and no copies of W303-1A-derived SNPs within the boxed region. Thus, this region represents a 3:1 conversion tract. In Figure 7C, in the boxed regions of both the red and white sectors, the hybridization signals indicate that there are two copies of YJM789-derived SNPs and no copies of W303-1A-derived SNPs, as expected for a 4:0 conversion event. Among spontaneous crossovers, an additional pattern of conversion associated with crossovers is a hybrid 4:0/3:1 tract. This pattern is consistent with a G_1-associated DSB in which the repair of the two broken sister chromatids results in gene conversion tracts of different lengths (St Charles and Petes 2013).

In our previous analyses of spontaneous RCOs in a wild-type strain, most events were consistent with DSB formation in G_1. The ratio of the numbers of 4:0 (or 4:0/3:1 hybrid) tracts to 3:1 tracts to no detectable tracts was 90:28:20 (St Charles and Petes 2013). The corresponding ratios for the *rnh201*Δ and *rnh1*Δ *rnh201*Δ strains in the current analysis were 7:2:12 and 2:8:4, respectively. These data are summarized in Figure 8. Though the numbers of sectored colonies analyzed in the current study were relatively small, the distribution of events in the *rnh201*Δ or *rnh1*Δ *rnh201*Δ diploid was significantly different from that in the wild-type strain ($P < 0.001$, Fisher exact test). For the *rnh201*Δ single mutant, the difference was driven by an increase in events with no detectable conversion tracts; when just 3:1 and 4:0 events were considered, there was no significant difference from wild type ($P = 1$). For the *rnh1*Δ *rnh201*Δ double mutant, however, the difference reflected a strong shift from predominantly G_1 events in wild type to predominantly S/G_2 events in the double mutant. The distribution of events in the *rnh201*Δ single mutant also was significantly different from that in the *rnh1*Δ *rnh201*Δ double mutant ($P = 0.016$, Fisher exact test).

In the wild-type strain, only 20 of 138 (14%) RCOs had no detectable gene-conversion tract, whereas in the *rnh201*Δ strain 12 of 21 crossovers (57%) had no detectable tract ($P < 0.001$ by Fisher exact test). The simplest explanation

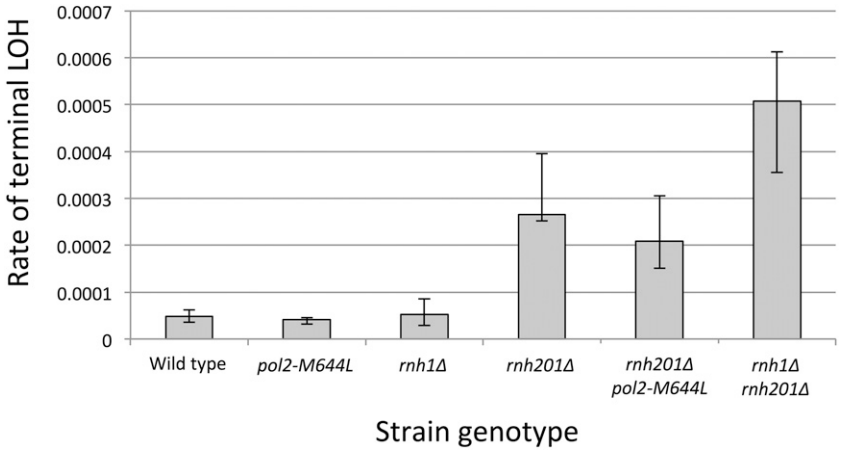

Figure 5 Rate of terminal LOH on chromosome IV in RNase H mutant diploids. LOH was assessed by measuring the rate of 5-FOA resistance, which corresponds to loss of the *URA3* marker near the end of chromosome IV. For wild-type, *pol2-M644L*, *rnh1*Δ, *rnh201*Δ, *rnh201*Δ *pol2M644L*, and *rnh1*Δ *rnh201*Δ diploids, 17, 23, 16, 17, 25, and 19 independent cultures were used to derive the rates of instability, respectively.

of this difference is that the conversion tracts in the *rnh201*Δ strain are shorter than in the wild-type strain and, therefore, less likely to include a heterozygous SNP. Indeed, direct measurements of the conversion tract lengths confirmed this expectation, with tracts being significantly shorter in the *rnh201*Δ mutant than in wild type ($P = 0.03$, Wilcoxon rank sum test); conversion tract lengths in the *rnh1*Δ *rnh201*Δ strain were not significantly different from wild type ($P = 0.63$). The median tract lengths in the wild-type, *rnh201*Δ, and *rnh1*Δ *rnh201*Δ (95% confidence limits in parentheses) strains were 10.6 kb (8.2–13.6 kb; (St Charles and Petes 2013), 4.8 kb (1.7–17.1 kb), and 9.3 kb (2.2–19 kb), respectively. It should be emphasized that these conversion tracts are all associated with a crossover. The conversion tract lengths from the subcultured strains in Table S5 represent a mixture of conversion events that are associated and unassociated with crossovers.

Depictions of each class of sector identified in the current analysis are shown in Figure S7 and the coordinates of the breakpoints in each event are listed in Table S8. We analyzed events on chromosome IV in *rnh201*Δ and *rnh1*Δ *rnh201*Δ mutants for enrichment of various genetic elements using the same procedure as used for the subcultured strains. None of the genomic elements listed in Table S10 were significantly over- or underrepresented at recombination breakpoints on the right arm of chromosome IV (Table S12).

Discussion

Previous work has shown that strains lacking RNase H1 and/or RNase H2 have elevated levels of mitotic recombination and chromosome loss. RNase H2-defective haploids, for example, have elevated gene conversion between closely linked repeats (Potenski *et al.* 2014) and increased loss of markers flanked by direct repeats (Ii *et al.* 2011). Although loss of either RNase H1 or RNase H2 promoted loss of an artificial chromosome, loss of both enzymes was required to elevate LOH on chromosome III in a diploid background (Wahba *et al.* 2011). Our work extends analyses of instability to a genome-wide scale in diploid yeast strains that are partially (*rnh1*Δ and *rnh201*Δ single mutants) or completely (*rnh1*Δ *rnh201*Δ double mutant) defective in RNase H activity and uses microarrays to provide a high-resolution map of recombination events resulting in LOH. Significantly, in each of three assays used, LOH was elevated in the *rnh201*Δ, but not the *rnh1*Δ single mutant, and was further elevated in the *rnh1*Δ *rnh201*Δ double mutant. We additionally used a mutant DNA polymerase that lowers the direct incorporation of rNMPs into genomic DNA, allowing us to assess the relative contributions of R-loops *vs.* rNMPs to LOH in the *rnh201*Δ background. The discussion below focuses on three related issues: (1) the nature of the recombination-initiating lesion in strains lacking RNase H activity, (2) the timing of formation of the recombinogenic lesion, and (3) factors that regulate the distribution of recombination events associated with loss of RNase H activity.

Nature of the recombinogenic lesions in strains lacking RNH1 and/or RNH201

The DNA alterations that lead to elevated recombination in strains lacking *RNH1* or *RNH201* are likely rNMPs embedded in DNA (expected to accumulate in *rnh201*Δ strains) and/or R-loops (expected to accumulate in *rnh1*Δ and *rnh201*Δ strains). In the case of loss of RNase H2, the elevated level of intrachromosomal recombination between repeats is dependent on Topoisomerase I (Top1) (Potenski *et al.* 2014). Further genetic and biochemical studies suggest that Top1-mediated cleavage at rNMPs is followed by the sequential action of Srs2 and Exo1, which produces a single-strand gap (Potenski *et al.* 2014). Subsequent replication of a gap-containing chromosome would be expected to produce a broken, presumably recombinogenic, chromatid. That R-loop accumulation results in a hyperrec phenotype has been shown using mutants defective in transcript processing and/or in the removal of R-loops (Aguilera and Garcia-Muse 2012; Hamperl and Cimprich 2014). The corresponding recombinogenic DNA lesion could reflect either nicking of the unpaired DNA strand within an R-loop or conflicts between the replication fork and an R-loop. Either of these mechanisms would likely result in single broken sister

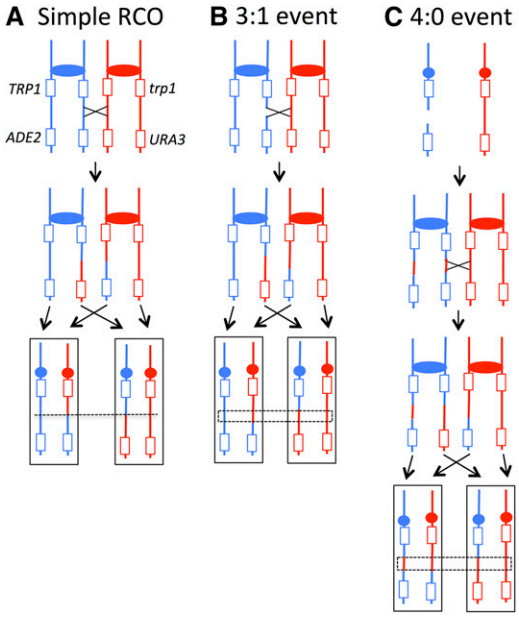

Figure 6 Classes of red/white sectors resulting from RCO. As in Figure 3, blue and red lines indicate YJM789- and W303-1A-derived chromatids, respectively. Three common types of crossovers that result in a red Ura$^+$ and a white Ura$^-$ sector can be distinguished by microarray analysis. (A) If a RCO is not associated with a gene conversion, the transition between heterozygous and homozygous SNPs occurs at the same position in the two sectors. Such crossovers provide no information about the likely timing of the recombinogenic DNA lesion. (B) A DSB formed in S or G_2 on the YJM789-derived chromatid is associated with the nonreciprocal transfer of information from the W303-1A-derived chromatid, producing a 3:1 gene conversion event. In the dotted box, there are three chromosomes with W303-1A-derived sequences and only one chromosome with YJM789-derived sequences. By microarray analysis, the red sector loses heterozygosity at a more centromere-proximal location than the white sector. (C) A DSB occurs on the YJM789-derived homolog in G_1, and the broken molecule is replicated to produce two sister chromatids that are broken at the same position. Repair of the two broken DNA molecules produces a 4:0 conversion tract as indicated by the dotted lines. Repair of one these breaks is associated with a crossover, producing the red/white sectored colony.

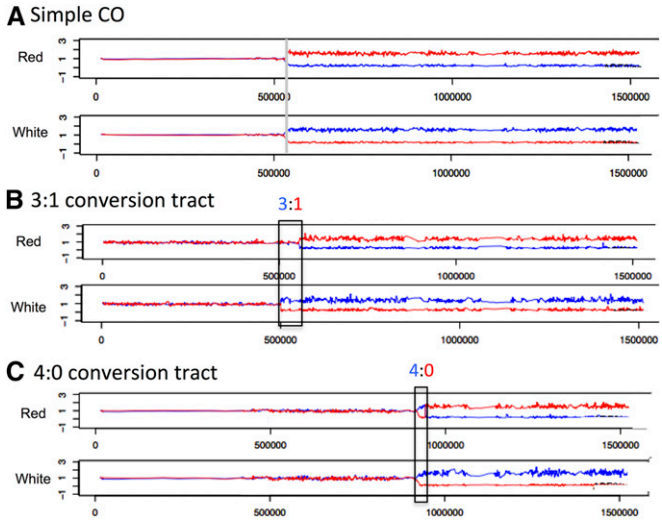

Figure 7 Examples of microarrays showing simple crossovers, crossovers with 3:1 conversion tracts, and crossovers with 4:0 conversion tracts. The red and white sectors of sectored colonies were examined by chromosome IV-specific microarrays. The hybridization ratio of DNA derived from the sectors relative to heterozygous control DNA is shown on the *y*-axis; the red and blue lines represent hybridization to W303-1A-specific or YJM789-specific oliognucleotides, respectively. The *x*-axis shows the SGD coordinates on the right arm of chromosome IV. (A) In this sectored colony (KO135_5_2R and KO135_5_2W in Table S8), the red and white sectors have the same point of transition between heterozygous and homozygous markers (near SGD coordinate 530 kb), indicative of a simple crossover. (B) In this sectored colony (KO132_31_17R and KO132_31_17W in Table S8), the red sector has a transition point at about coordinate 500 kb, and the white sector has a transition near coordinate 555 kb. Thus, this sectored colony has a large (~55 kb) 3:1 conversion tract associated with the crossover. In the boxed region, the red sector has one copy each of W303-1A- and YJM789-derived SNPs. In this region, the white sector has two copies of YJM789-derived SNPs and no copies of W303-1A-derived SNPs. (C) In this sectored colony (KO135_5_5R and KO135_5_5W in Table S8), there is a region of ~20 kb that is homozygous for the YJM789-derived SNPs in both red and white sectors, consistent with a 4:0 conversion.

chromatid in S phase. It is widely assumed that most, if not all, R-loops are redundantly processed by RNase H1 and RNase H2.

In our experiments, we did not detect a hyperrec phenotype in *rnh1*Δ single mutants, but consistently observed a three- to fivefold increase in LOH in *rnh201*Δ strains. In all three assays, instability in the *rnh1*Δ *rnh201*Δ strains was elevated relative to that in the *rnh201*Δ single-mutant strains. One interpretation of the lack of a hyperrec phenotype in the *rnh1*Δ mutant and the substantial hyperrec phenotype in the *rnh201*Δ mutant is that stimulation of recombination upon loss of RNase H2 is solely a consequence of misincorporated rNMPs. Two arguments suggest that this extreme hypothesis is not correct. First, since the hyperrec phenotype is stronger in the *rnh1 rnh201* double mutant than in the *rnh1* mutant, and RNase H1 has no activity on single ribonucleotides, the hyperrec phenotype of the *rnh201* strain must reflect, at least in part, some other type of lesion than single ribonucleotides. Second, if all of the recombination events in the *rnh201*Δ mutant reflect persistent rNMPs in DNA, then the hyperrec phenotype should be substantially reduced in an *rnh201*Δ *pol2-M644L* strain. Prior studies have demonstrated that the presence of the *pol2-M644L* allele reduces the level of rNMPs in genomic DNA ~70% relative to a strain with wild-type Pol 2 activity (Nick McElhinny *et al.* 2010a). Although we did observe a small reduction in LOH in the *rnh201*Δ *pol2-M644L* strain relative to the *rnh201*Δ strain in one of our LOH assays, the reduction was only ~20%. In a similar assay, the rate of LOH in an *rnh201*Δ strain was reduced less than twofold by the *pol2-M644L* mutation (Conover *et al.* 2015).

Based on subtle effect of the *pol2-M644L* mutation, we suggest that most of the LOH in the *rnh201*Δ mutant reflects persistent R-loops rather than persistent rNMPs. Our data indicate that RNase H2 can remove most, if not all, of the R-loops that accumulate in the absence of RNase H1, but that RNase H1 can remove only a relatively small fraction of the

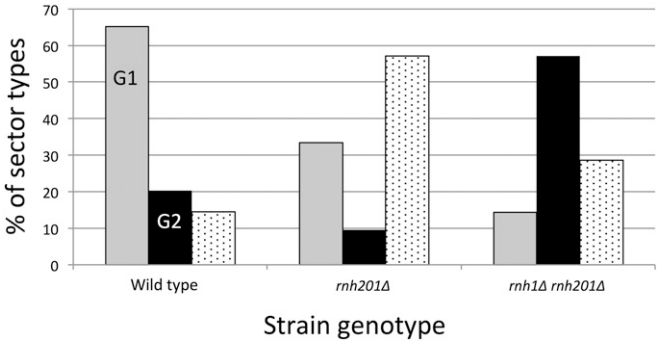

Figure 8 Distribution of sector types by cell cycle stage in which the initiating lesion arose. Conversion tracts with a 4:0 region are indicative of a G_1-associated DSB (gray bars), and 3:1 tracts indicate a G_2-/S-phase DSB (black bars). The stippled bars show simple CO events in which the timing of the recombinogenic lesion cannot be inferred.

R-loops that accumulate in the *rnh201*Δ mutant. Why this particular division of labor might not have been evident in prior studies could reflect the monitoring of instability in much smaller genetic intervals and/or the examination of RNase H activity only under conditions of pathological R-loop accumulation. The ability of RNase H2 to fully compensate for loss of RNase H1 activity may also be related to the induction of RNase H2 in strains that lack RNase H1 (Arudchandran *et al.* 2000).

Is there any role of misincorporated rNMPs in stimulating LOH in diploids? Several types of data suggest that there is. First, as discussed above, Potenski *et al.* (2014) defined a pathway in which Top1 acts on rNMPs to produce recombinogenic lesions; a similar pathway produces mutagenic DNA lesions (Kim *et al.* 2011). Second, we observed a small reduction in LOH in the *rnh201*Δ *pol2-M644L* strain relative to the *rnh201*Δ strain; this reduction was statistically significant, however, only for the assay measuring the rate of LOH on chromosome IV (Figure 5). Altogether, our data support the conclusion that recombination events in strains lacking RNase H1 and H2 are primarily a consequence of R-loop formation, with misincorporated rNMPs playing only a minor role. This conclusion, however, is based on the assumption that the reduced incorporation of ribonucleotides resulting from the *pol2-M644L*-encoded DNA polymerase is not affected by the *rnh201* mutation or other aspects of the genetic background in the *rnh201 pol2-M644L* diploid. In addition, we assume that the *rnh201 pol2-M644L* strain does not have an alteration (for example, a significantly slower S phase) that affects the likelihood that the repair template is the sister chromatid rather than the homolog.

It should also be noted that triple mutant combination of *rnh1*Δ *rnh201*Δ *pol2-M644G* is synthetically lethal, whereas the double mutant *rnh201*Δ *pol2-M644G* is viable (Lazzaro *et al.* 2012). Since the *pol2-M644G* allele encodes a form of Pol ε that incorporates increased levels of ribonucleotides (Nick McElhinny *et al.* 2010a), this result was interpreted as indicating a possible role of RNase H1 in the removal of ribonucleotides (Lazzaro *et al.* 2012), although other interpretations of the synthetic lethality are possible.

Cell-cycle timing of recombinogenic lesions in strains lacking RNase H1 and/or H2

Most of the proposed models for the recombinogenic effects of R-loops or misincorporated rNMPs predict a DNA lesion that leads to one broken chromatid in an S- or G_2-phase cell. Such a lesion could be repaired by sister-chromatid exchange (no observable LOH) or by an interaction with the homolog that is expected to result in an associated 3:1 conversion event. In a wild-type strain, only ~30% of sectored colonies have the 3:1 pattern, indicating that most LOH is initiated in G_1 (St Charles and Petes 2013). Most of the events in the *rnh1*Δ *rnh201*Δ double-mutant strain had the S/G_2 pattern, however, with 8 of the 10 sectored colonies having 3:1 conversion tracts. This result is consistent with the recombinogenic lesion resulting from an interaction of the replication fork with an R-loop. In contrast, in the *rnh201*Δ strain, seven of the nine events had 4:0 or 3:1/4:0 conversion tracts indicative of a G_1-initiated event, and thus were similar to wild type. One interpretation of this result is that RNase H2 might be specialized to remove R-loops that give rise to recombinogenic lesions outside of S phase. Indeed, immunological data suggest that ~30% of pathological R-loops exert their recombinogenic effect outside of S phase (Wahba *et al.* 2011). The transcription of *RNH201* is elevated prior to S phase and continues at a high level in S (Pramila *et al.* 2006); the transcription of *RNH1* is not periodic (Pramila *et al.* 2006). Alternatively, this result could reflect the production of a double-stranded DNA break reflecting the processing of rNMPs that are close (<10 bp) together on opposite strands of duplex DNA.

Distribution of recombination events associated with loss of RNase H activity

With the possible exception of the ribosomal RNA genes, no strong hotspots for LOH were observed in the subcultured *rnh201*Δ, *rnh201*Δ *pol2-M644L*, or *rnh1*Δ *rnh201*Δ strains (Figure S4, Figure S5, and Figure S6). In particular, we found no correlation of LOH breakpoints with a number of chromosome elements expected to be associated with R-loop formation such as highly transcribed genes, intron-containing genes, or G4-forming motifs. We also observed no strong correlation between recombination breakpoints in subcultured mutant strains and regions of R-loop accumulation. Finally, we found no association between regions of Rpb3p (a subunit of RNA polymerase II) accumulation and recombination breakpoints in the *rnh201*Δ, *rnh201*Δ *pol2-M644L*, or *rnh1*Δ *rnh201*Δ strains. These observations suggest that R-loop associated recombinogenic lesions are widely distributed throughout the yeast genome and/or that a number of different factors each contribute to the hyperrec phenotype of strains lacking RNase H activity.

Although several studies reported that Ty elements accumulated RNA:DNA hybrids in strains lacking RNase H activity (Chan *et al.* 2014; El Hage *et al.* 2014), we found no enrichment of Ty elements among LOH events. We did, however,

find that deletion and duplication events frequently resulted from homologous recombination between nonallelic Ty elements (Table S6). It is difficult to assess the significance of this observation since Ty elements are the primary type of large dispersed repeats in the yeast genome. In addition, nonallelic recombination between Ty elements is a common source of chromosome rearrangements in mutant yeast strains that do not accumulate R-loops (McCulley and Petes 2010; Song *et al.* 2014). Several other relevant factors should be mentioned. First, our experiments were performed in diploid strains with both *MAT***a** and *MAT*α information, and previous studies showed that Ty transcription is repressed in such diploids (Errede *et al.* 1980). Second, RNA:DNA hybrids associated with Ty elements primarily involve cDNA copies rather than the genomic elements (El Hage *et al.* 2014). Finally, it should be pointed out that our genetic assays are fundamentally different than the physical analysis of R-loop formation. To be detected by our LOH assays, the recombination event stimulated by R-loop formation must involve an interaction of the broken chromosome with the other homolog; breaks that are repaired by equal sister-chromatid recombination are genetically silent.

A significant enrichment of LOH events near the telomere was observed in the *rnh201*Δ strain. Subtelomeric regions encode a telomeric-repeat-containing RNA (TERRA) that accumulates in strains lacking RNase H2 (Yu *et al.* 2014). Strains with increased levels of telomeric RNA:DNA hybrids have elevated rates of telomere–telomere recombination (Yu *et al.* 2014). If elevated R-loops at the telomere cause a partial defect in telomere elongation by telomerase, there may be increased degradation of the ends of the chromosome, resulting in elevated levels of telomere-associated LOH. We note that Hackett and Greider (2003) previously showed an increase in terminal LOH in strains that had telomerase defects.

Summary

Our genome-wide analysis of instability associated with loss of RNase H in yeast shows that RNase H2 activity is much more important than RNase H1 activity in the maintenance of genome stability. However, strains that lack both RNase H1 and RNase H2 have qualitative and quantitative differences in genome stability relative to strains that lack only RNase H2. Our results suggest that R-loops contribute to most of the genetic instability of strains lacking RNase H activity, and that RNase H2 is uniquely able to process a subpopulation of R-loops. These results are relevant to human pathologies associated with defects in RNase H2 as well as the particular species of RNA:DNA hybrids that serve as the triggers for autoimmune disease. In the specific case of Aicardi-Goutiéres syndrome, recent work suggests that RNA:DNA hybrids are the likely immunogenic trigger of disease (Lim *et al.* 2015).

Acknowledgments

We thank J. L. Argueso and D. Koshland for communicating unpublished information and all members of the Petes and Jinks-Robertson labs for useful suggestions. The research was supported by National Institutes of Health grants GM24110 and GM52319 to T.D.P., and GM038464 and GM101690 to S.J.-R. K.O. was supported by a National Science Foundation graduate research fellowship (1106401).

Note added in proof: See Conover *et al.* 2015 (pp. 951–961) in this issue for a related work.

Literature Cited

Aguilera, A., and T. Garcia-Muse, 2012 R loops: from transcription byproducts to threats to genome stability. Mol. Cell 46: 115–124.

Altman, D. G., 1991 *Practical Statistics for Medical Research*, Chapman and Hall/CRC Press, Boca Raton, FL.

Arudchandran, A., S. Cerritelli, S. Narimatsu, M. Itaya, D. Y. Shin *et al.*, 2000 The absence of ribonuclease H1 or H2 alters the sensitivity of *Saccharomyces cerevisiae* to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. Genes Cells 5: 789–802.

Azvolinsky, A., P. G. Giresi, J. D. Lieb, and V. A. Zakian, 2009 Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. Mol. Cell 34: 722–734.

Cerritelli, S. M., and R. J. Crouch, 2009 Ribonuclease H: the enzymes in eukaryotes. FEBS J. 276: 1494–1505.

Chan, Y. A., M. J. Aristizabal, P. Y. Lu, Z. Luo, A. Hamza *et al.*, 2014 Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet. 10: e1004288.

Clausen, A. R., S. A. Lujan, A. B. Burkholder, C. D. Orebaugh, J. S. Williams *et al.*, 2015 Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. Nat. Struct. Mol. Biol. 22: 185–191.

Conover, H., and S. Lujan, M. Chapman, D. Cornelio, R. Sharif *et al.* 2015 Stimulation of chromosomal rearrangements by ribonucleotides. Genetics 947–957.

Crow, Y. J., A. Leitch, B. E. Hayward, A. Garner, R. Parmar *et al.*, 2006 Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutiéres syndrome and mimic congenital viral brain infection. Nat. Genet. 38: 910–916.

El Hage, A., S. Webb, A. Kerr, and D. Tollervey, 2014 Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet. 10: e1004716.

Errede, B., T. S. Cardillo, F. Sherman, E. Dubois, J. Deschamps *et al.*, 1980 Mating signals control expression of mutations resulting from insertion of a transposable repetitive element adjacent to diverse yeast genes. Cell 22: 427–436.

Fachinetti, D., R. Bermejo, A. Cocito, S. Minardi, Y. Katou *et al.*, 2010 Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. Mol. Cell 39: 595–605.

Günther, C., B. Kind, M. A. Reijns, N. Berndt, M. Martinez-Bueno *et al.*, 2015 Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. J. Clin. Invest. 125: 413–424.

Hackett, J. A., and C. W. Greider, 2003 End resection initiates genomic instability in the absence of telomerase. Mol. Cell. Biol. 23: 8450–8461.

Hamperl, S., and K. A. Cimprich, 2014 The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. DNA Repair (Amst.) 19: 84–94.

Hochberg, Y., and Y. Benjamini, 1990 More powerful procedures for multiple significance testing. Stat. Med. 9: 811–818.

Ii, M., T. Ii, L. I. Mironova, and S. J. Brill, 2011 Epistasis analysis between homologous recombination genes in *Saccharomyces*

cerevisiae identifies multiple repair pathways for Sgs1, Mus81-Mms4 and RNase H2. Mutat. Res. 714: 33–43.

Kim, N., S. Y. Huang, J. S. Williams, Y. C. Li, A. B. Clark *et al.*, 2011 Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. Science 332: 1561–1564.

Koh, K. D., S. Balachander, J. R. Hesselberth, and F. Storici, 2015 Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. Nat. Methods 12: 251–257.

Lazzaro, F., D. Novarina, F. Amara, D. L. Watt, J. E. Stone *et al.*, 2012 RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. Mol. Cell 45: 99–110.

Lea, D. E., and C. A. Coulson, 1949 The distribution of the numbers of mutants in bacterial populations. J. Genet. 49: 264–285.

Lee, P. S., and T. D. Petes, 2010 Mitotic gene conversion events induced in G1-synchronized yeast cells by gamma rays are similar to spontaneous conversion events. Proc. Natl. Acad. Sci. USA 107: 7383–7388.

Lim, Y. W., L. A. Sanz, X. Xu, S. R. Hartono, and F. Chédin, 2015 Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutières syndrome. eLife 4: doi: 10.7554/eLife.08007.

McCulley, J. L., and T. D. Petes, 2010 Chromosome rearrangements and aneuploidy in yeast strains lacking both Tel1p and Mec1p reflect deficiencies in two different mechanisms. Proc. Natl. Acad. Sci. USA 107: 11465–11470.

Mischo, H. E., B. Gomez-Gonzalez, P. Grzechnik, A. G. Rondon, W. Wei *et al.*, 2011 Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol. Cell 41: 21–32.

Nick McElhinny, S. A., D. Kumar, A. B. Clark, D. L. Watt, B. E. Watts *et al.*, 2010a Genome instability due to ribonucleotide incorporation into DNA. Nat. Chem. Biol. 6: 774–781.

Nick McElhinny, S. A., B. E. Watts, D. Kumar, D. L. Watt, E. B. Lundstrom *et al.*, 2010b Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. Proc. Natl. Acad. Sci. USA 107: 4949–4954.

Potenski, C. J., H. Niu, P. Sung, and H. L. Klein, 2014 Avoidance of ribonucleotide-induced mutations by RNase H2 and Srs2-Exo1 mechanisms. Nature 511: 251–254.

Prado, F., and A. Aguilera, 2005 Impairment of replication fork progression mediates RNA polII transcription-associated recombination. EMBO J. 24: 1267–1276.

Pramila, T., W. Wu, S. Miles, W. S. Noble, and L. L. Breeden, 2006 The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle. Genes Dev. 20: 2266–2278.

Reijns, M. A., H. Kemp, J. Ding, S. M. De Proce, A. P. Jackson *et al.*, 2015 Lagging-strand replication shapes the mutational landscape of the genome. Nature 518: 502–506.

Song, W., M. Dominska, P. W. Greenwell, and T. D. Petes, 2014 Genome-wide high-resolution mapping of chromosome fragile sites in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 111: E2210–E2218.

St Charles, J., and T. D. Petes, 2013 High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1 Mb arm of yeast chromosome IV. PLoS Genet. 9: e1003434.

St Charles, J., E. Hazkani-Covo, Y. Yin, S. L. Andersen, F. S. Dietrich *et al.*, 2012 High-resolution genome-wide analysis of irradiated (UV and gamma-rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events. Genetics 190: 1267–1284.

Wahba, L., J. D. Amon, D. Koshland, and M. Vuica-Ross, 2011 RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol. Cell 44: 978–988.

Williams, J. S., and T. A. Kunkel, 2014 Ribonucleotides in DNA: origins, repair and consequences. DNA Repair (Amst.) 19: 27–37.

Williams, J. S., D. J. Smith, L. Marjavaara, S. A. Lujan, A. Chabes *et al.*, 2013 Topoisomerase 1-mediated removal of ribonucleotides from nascent leading-strand DNA. Mol. Cell 49: 1010–1015.

Yin, Y., and T. D. Petes, 2013 Genome-wide high-resolution mapping of UV-induced mitotic recombination events in *Saccharomyces cerevisiae*. PLoS Genet. 9: e1003894.

Yu, T. Y., Y. W. Kao, and J. J. Lin, 2014 Telomeric transcripts stimulate telomere recombination to suppress senescence in cells lacking telomerase. Proc. Natl. Acad. Sci. USA 111: 3377–3382.

Communicating editor: N. Hollingsworth

GENETICS

Supporting Information
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.182725/-/DC1

Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity

Karen O'Connell, Sue Jinks-Robertson, and Thomas D. Petes

DOI: 10.1534/genetics.115.182725

File S1. Expanded Materials and Methods

Strain construction

Most of the details of the strain constructions are in Table S1. As in previous studies (for example, Lee *et al.*, 2009), we used derivatives of the sequence-diverged haploids W303-1A and YJM789. These haploids with various genetic alterations were crossed to generate the diploids used to assay genetic instability. Most haploid strains were constructed by transformation with PCR-generated DNA fragments or by sporulating nearly-isogenic diploids. The genotypes of spores for auxotrophic markers were determined by replica-plating spore-derived colonies to omission media. The replacements of genes with drug-resistance markers were confirmed by PCR analysis as described in Tables S2 and S3. Mating type was determined by PCR with primers MATaF, MATalphaF, and MATR (Table S3). *MAT***a** and *MAT*α loci were associated with 500 and 400 bp fragments, respectively.

Measurements of rates of genetic instability induced by loss of RNase H

We used three methods to examine the rates of instability in strains with mutations affecting RNase H activity. We first measured the frequency of genomic alterations in sub-cultured diploid strains of the following genotypes: wild-type, *rnh1*Δ, *rnh201*Δ, *rnh1*Δ *rnh201*Δ, *pol2-M644L*, and *rnh201*Δ *pol2-M644L*. All diploids were generated by crossing haploids isogenic with W303-1A and YJM789. Two independently-derived isolates from each strain were streaked with a toothpick to single colony density on rich growth medium (YPD) for the first subculture. For the second subculture, five-ten colonies from each isolate were then re-streaked to YPD. One colony derived from each of these ten to twenty colonies was then re-streaked again. For the mutant backgrounds, this procedure was repeated twenty times. For the wild-type, only ten sub-culturings

were performed. All sub-culturing experiments were done at 30^0 C. Following sub-culturing, the passaged strains were examined by whole-genome microarrays as described below.

The second method of measuring genome stability was to monitor the frequency of formation of red/white sectored colonies in strains in which the *ADE2* gene was inserted near the right telomere of one copy of chromosome IV and the *K. lactis URA3* was inserted at the allelic position on the other copy (Fig. 3). The homologs with the *ADE2* and *URA3* genes had wild-type and mutant alleles of the centromere-linked *TRP1* gene, respectively. Experiments were initiated using colonies formed on YPD plates. For each genotype, we used one colony from two independently-constructed diploids. Each colony was suspended in water, and diluted to a concentration that resulted in about 1000 cells per plate on the diagnostic medium (SD-arginine with 10 micrograms/ml adenine). After three days of growth, we scored plates for red/white sectors using a dissecting microscope. Cells from each sector were then re-streaked to YPD plates and, after two days of growth, the resulting colonies were replica-plated to media lacking uracil or tryptophan. If all of the colonies purified from the white sector were Ura^-, the sectored colony was classified as resulting from a reciprocal crossover (Fig. 3A). If all colonies derived from the white sector were Ura^+, the strain was classified as resulting from BIR (Fig. 3B). White sectors that had mixtures of Ura^+ and Ura^- colonies were not used in our analysis; such sectored colonies could reflect events that occurred subsequent to the first division. Sectored colonies could also result from chromosome loss. In such colonies, the red sector would be Trp^-. Of 173 red/white sectored colonies, only one example of chromosome loss was observed.

The assay of genome instability based on red/white colony formation has the unfortunate characteristic of being non-selective. For the third assay, we selected for loss of the heterozygous *URA3* gene located near the right telomere by plating cells on

medium containing 5-fluoro-orotate (5-FOA). For this assay, we suspended colonies of each strain in 1 ml of water, and plated about 100 microliters of each undiluted suspension on medium containing 5-FOA (1 mg/ml), and a dilution of the suspension on non-selective medium (SD-complete medium). Between 15 and 25 colonies were examined for each genotype. From measurements of the number of 5-FOA^R colonies and the total number of cells in each colony/culture, we calculated the rate of 5-FOA^R using the method of the median (Lea and Coulson, 1949). To obtain the 95% confidence intervals for the rate estimates, we used Table B11 in Altman (1990).

Microarray analysis

DNA samples for microarray analysis were prepared by methods similar to those described in St. Charles *et al.* (2012). In brief, yeast cells were grown to stationary phase in liquid YPD cultures (5-15 ml). Cells were harvested by centrifugation, and the resulting cell pellet was resuspended in about 500 microliters of 42^0 molten agarose (0.5% low-melt agarose in 100 mM EDTA); 20 microliters of a 25 mg/ml solution of Zymolyase was then added. This mixture was distributed among about seven plug molds, each containing about 100 microliters. The samples were allowed to solidify at 4°C for 30 minutes. After solidifying, the plugs were suspended in 1ml of 10mM Tris/500mM EDTA (TE buffer), and incubated in a 15 ml tube at 37°C for at least 10 hours. 100 microliters of a 5% sarcosyl, 5 mg/ml proteinase-K in 500mM of EDTA (pH7.5) solution was then added to each sample, and the samples were incubated at 50° C for at least 12 hours. Each plug was then washed twice with TE buffer. The second incubation was performed with shaking at 4° C for at least 12 hours. After 12 hours, we did a third wash at 4^0 with TE buffer.

DNA was then extracted from four plugs of each sample using methods described in St. Charles *et al.* (2012). The samples were sonicated to yield DNA fragments of about 250 bp. The samples derived from each plug were pooled for labeling with fluorescent dyes. Each sample had about 100 micrograms/ml of DNA, and about 10 microliters was used in the labeling reactions.

For our method of analysis, the hybridization of DNA derived from experimental strains with LOH events was performed in competition with DNA from control strains heterozygous for all SNPs. The experimental strains were labeled with Cy3-dUTP, whereas the control strain was labeled with Cy5-dUTP (details in St. Charles *et al.*, 2012). The labeled nucleotides were provided as part of the Invitrogen Bioprime Array CGH Genome Labeling Module. The control and experimental labeled samples were combined, and hybridized to the microarrays. Two types of custom-made Agilent microarrays were used, one to analyze LOH events throughout the genome (St. Charles *et al.*, 2012) and one to examine LOH events on the right arm of chromosome IV (St. Charles and Petes, 2013). The sequences and locations of oligonucleotides on the whole-genome array are in Table S5 (St. Charles *et al.*, 2012), and the sequences and locations of oligonucleotides on the chromosome IV-specific array are in Table S9 (St. Charles and Petes, 2013).

After hybridization (conditions described in St. Charles *et al.*, 2012), the arrays were scanned using the GenePix scanner and GenePix Pro 6.1 software. A GenePix Results (.gpr) file was generated for each sample using the Batch Analysis feature in Gene Pix Pro 6.1. This file contains a “ratio of medians (635 nm/532 nm)” for each oligonucleotide represented on the microarray. This ratio reflects the fluorescence of the Cy5-labeled experimental sample relative to the Cy3-labeled control. Each SNP analyzed was represented by at least four 25-base oligonucleotides, two identical to the Watson and Crick strands containing the W303-1A allele and two identical to the Watson and Crick

strands of the YJM789 allele. We used programs written in Perl and R to automate the analysis and to plot hybridization levels throughout the genome (programs available on request). The resulting plots were done at two levels of resolution. Low-resolution plots depicted hybridization ratios that are the moving average of ten adjacent SNPs, whereas high-resolution plots show the ratio of medians at each individual SNP. We eliminated from the analysis any oligonucleotides that were "flagged" by the GenePix Pro 6.1 software or that had a level of fluorescence that was in the bottom 5% of intensities in both the 635 nm and 532 nm channels.

Following normalization, the hybridization ratios of heterozygous SNPs for the experimental samples were about 1 to both the W303-1A- and YJM789-related SNPs. For LOH regions in which W303-1A-related SNPs were homozygous, the ratio of hybridization to these SNPs was about 1.5 and the ratio of hybridization to YJM789-related SNPs was about 0.2. For LOH regions in which YJM789-related SNPs became homozygous, these ratios were reversed.

Microarray slides were re-used about six times by stripping the slides of the labeled DNA. Microarray slides containing DNA probes were stripped by placing them in stripping buffer (10 mM potassium phosphate, pH 6.6), and slowly heating to them to the boiling point over about one hour. After rinsing in water, the slides were stored in a nitrogen-containing cabinet. The slides that cover the microarray slides (gasket slides) were boiled in the stripping buffer for 40 minutes, rinsed in water, and dried by centrifugation. Following stripping, we usually allowed the slides to dry for two days before they were used again.

Associating genomic elements with LOH transitions in sub-cultured strains

One of the main goals of this research was to find out whether the LOH events resulting from loss of RNase H were enriched at the locations of specific genomic

elements. The rationale for our analysis is that the breakpoints associated with LOH events should be located near the site of the recombinogenic lesion. In our analysis, we examined only events in which at least two or more adjacent SNPs underwent LOH. To determine the likely "window" containing the recombination initiation site, we used the same procedure employed in our previous studies (for example, St. Charles and Petes, 2013). For interstitial LOH events, we used an association window that included all sequences located between the heterozygous SNPs that most closely flanked the LOH region. For terminal LOH events, the association window was 10 kb centromere-proximal and 10 kb centromere-distal from the homozygous SNP that was closest to the LOH event. Some events were found in most or all of the sub-cultured strains derived from a single isolate. Since these events (marked "redundant" in Table S5) were likely generated in the isolate before sub-culturing, we included redundant events in each category as single events.

We next determined whether specific genetic elements were over-represented in the association windows of different mutant strains. This analysis involved multiple steps. First, for each genotype, we summed the number of bases in the association windows over all of the individual sub-cultured isolates. For example, for the nineteen sub-cultured isolates of the *rnh201* strain about 1.35 Mb were included in the association windows (Table S5). The yeast nuclear genome as annotated in SGD, which includes only two of the approximately 150 rRNA gene repeats, is about 12.1 Mb. As discussed in the legend to Table S10, our microarrays cover about 11.6 Mb of the genome, since these arrays do not include the repetitive sub-telomeric sequences. Thus, the total amount of DNA represented on the arrays for 19 isolates is about 220.4 Mb. The amount of genomic DNA that is not present in the association windows is, therefore, 220.4 Mb – 1.35 Mb or about 219 Mb. Second, we determined the total number of specific genomic elements represented on the array. For example, there are 352 ARS elements in the

genome and 317 ARS elements represented on the array (Table S10). If these elements are placed randomly with respect to the association windows, we expect 37 ARS elements within the association windows: 317 x (1.35 Mb/220.4 Mb) x 19. The expected number of ARS elements located outside of the association windows is 5986: (317 x 19) – 37. We then counted the number of ARS elements within the association windows, determining that there were 31; the observed number of these elements outside of the association windows was 5992. Finally, we compared the observed and expected numbers by Chi-square analysis (Table S11), finding a p value (0.362) that indicates no significant association between LOH breakpoints and ARS elements. We repeated this analysis with twenty other genomic features (described below). After correction of the p values for multiple comparisons (Benjamini and Hochberg, 1995), none of these values were significant.

The numbers and locations of each genomic element tested were assembled from a variety of sources. Ty elements, solo LTR elements, centromeres, intron-containing genes, ARS elements, and tRNA genes were extracted from the S288c reference genome using the YeastMine tool on SGD (Engel *et al.* 2013, genome version R64-1-1; http://www.yeastgenome.org/help/video-tutorials/yeastmine). We also used YeastMine to determine the locations of the genes that were among the top 5% in length ("long gene" category in Tables S10 and S11). The same tool was used to identify the genes with the highest (top 5%) and lowest (bottom 5%) rates of transcription. From the sequence of the ORFs, we calculated the percentage of G bases on the non-transcribed strand. 41 ORFs with ≥ 29% G were identified and used in the association analysis. In addition, we identified 115 ORFs that had a GC-content ≥ 50%; these genes were also used in our analysis (Table S11).

Most of the other references for the locations of various genomic elements are in Table S9. Regions with converging replication forks (TER sites) were described in Table

S2 of Fachinetti *et al*. (2010). In the same paper, binding sites for the Rbp3p subunit of RNA polymerase II were mapped by chromatin immunoprecipitation followed by microarray analysis. We downloaded these data (GSM409326 on GEO, GSM409326_Rpb3_signal.bar.gz) and converted them to a .txt file. All sites with a normalized log2 value less than 0.4 were eliminated from analysis. Adjacent sites less than 1 kb apart were collapsed into single intervals, and the signal was averaged over all collapsed sites. There were 933 such intervals. As sorted by the hybridization values, we used the top 10% (93) of the intervals for our association analysis. 58 of the 71 genomic TER sites were associated with Rpb3p binding (Table S4; Fachinetti *et al.*, 2010). These sites were designated "TER sites related to high transcription" in Tables S9-S12. We also examined the association of LOH breakpoints with sites enriched for the binding of Rrm3p, a helicase involved in promoting replication through certain hard-to-replicate sequences; the map locations of these sites are in Supplemental Table 7 of Azvolinsky *et al.* (2009).

Palindromic sequences greater than 16 bp were examined using data from Lisnic *et al.* (2005), and sequences likely to form G4 quadruplex structures were obtained from Dataset S1 of Capra *et al*. (2010). Hershman *et al.* (2008) examined differential expression of genes by *S. cerevisiae* in response to N-methyl mesoporphyrin IX (NMM), a drug that stabilizes G4 quadruplexes *in vitro*. We examined the association of genes whose transcription was significantly ($p<0.001$) altered by the drug (Supplementary Table 5 of Hershman *et al.*, 2008) with the LOH events. We also examined associations with genomic regions with high levels of Elc1p, a protein involved in resolving conflicts between converging transcripts (Hobson *et al.*, 2012). The top 10% of Elc1p-binding sites (Table S1 of Hobson *et al*., 2012) were used to look for associations. The locations of RNA/DNA hybrids in *rnh1Δ rnh201Δ* strains have been recently mapped (Chan *et al.*, 2014). We examined the association of those sites that were at least ten-fold enriched

(Dataset S7 of Chan *et al.*, 2014) with our LOH data. Lastly, based on observations of a non-random association of R-loops and poly A tracts (Doug Koshland, University of California, Berkeley), we identified all uninterrupted poly A or poly T tracts that were at least 25 bases in length. We examined the association of these 41 tracts with LOH events. As described in the main text, after correction for multiple comparisons, none of the genomic elements that we examined were significantly associated with the LOH events in sub-cultured strains.

In our analysis, any elements that are within the association windows or that span the association windows are included in our analysis. For most of the genomic elements examined, the size of the element was small, less than 10% of the average size of the association window. For four of the elements (Ty elements, TER sites, TER sites associated with high levels of transcription, and "Long Genes"), however, the size of the element was greater than 10% of the size of the association window. For these comparisons, we expanded all association windows by an amount equivalent to the average size of the element. For example, when we examined the associations between Ty elements and LOH events, the association windows were expanded by 6 kb, the size of a Ty element.

Associating genomic elements with LOH transitions in sectored colonies

In sectored colonies (reflecting crossovers on the right arm of chromosome IV), the borders of the association window were the coordinates of the heterozygous SNP closest to the most centromere-proximal LOH transition and the homozygous SNP closest to the most centromere-distal LOH transition (Table S8). All association windows were used in our analysis. The right arm of chromosome IV represents about 1.4 Mb, and the number of genomic elements on the right arm of IV are given in Table S10. Our methods of calculating significant associations between genomic elements and LOH

events are chromosome IV were analogous to those described for the sub-cultured strains. We performed microarray analysis on sectored colonies of only two genotypes: *rnh201*Δ and *rnh1*Δ *rnh201*Δ. No significant associations were found between LOH events in these strains and any of the tested genomic elements (Table S12).

Regions of apparent terminal duplications/deletions at repetitive sub-telomeric regions

By microarray analysis, regions of LOH are unambiguous since the hybridization signals for one set of allelic SNPs increases for the same genomic region in which the other set of allelic SNPs decreases. From previous studies (Y. Yin and T. Petes, unpublished observations), we have found a small number of apparent terminal duplications and deletions that likely reflect LOH events on non-homologous chromosomes with shared sub-telomeric sequences. In the current study, we observed several such events among the sub-cultured strains as described below.

KO_244_1_XX_D (*rnh201*Δ *pol2-M644L*). In this isolate, we observe a terminal LOH event (YJM789-derived SNPs becoming homozygous) on chromosome X (transition coordinates 708414-728414). This strain also has a terminal deletion on the left arm of chromosome IV (transition coordinates 15561-18870), resulting in loss of W303-1A-derived sequences. In the sequence of S288c (nearly isogenic with W303-1A), we found that the chromosome X sequences between 730-742 kb are almost identical to the region 3-15 kb on chromosome IV. For example, the oligonucleotide 5211 (Table S3 in St. Charles *et al.,* 2012), near the right telomere of IV is repeated near the right telomere of X. Thus, an LOH event causing loss of W303-1A-derived SNPs from chromosome X will appear as a reduced signal of hybridization to W303-1A-specific SNPs near the right telomere of chromosome IV. It is unclear whether the YJM789-derived copy of chromosome X has the same duplication as the W303-1A-derived homolog. When sequences from a portion of the repeated region on chromosome IV (10 kb to 12 kb) are

used in a BLAST search of the YJM789 database in SGD, the sequences match to chromosome IV contigs without matching to chromosome X contigs. Finally, we note that LOH events that involve the left end of IV will not have a detectable effect on the microarray pattern observed on the right end of chromosome X because the most centromere-distal oligonucleotide on the array is at position 727 kb which is centromere-proximal to the repeated sequences.

KO_5_6_E (*rnh1*Δ *rnh201*Δ); KO_5_9_H (*rnh1*Δ *rnh201*Δ); KO_75_2_XX_H (*rnh201*Δ). These three strains had an apparent terminal deletion on chromosome VI, resulting in loss of YJM789-derived sequences. The coordinates for the deletions were similar in all three isolates beginning near coordinate 30,000 kb and proceeding to the telomere. All three strains also had terminal LOH events on chromosome X, resulting in loss of YJM789-derived SNPs and duplication of W303-1A-derived SNPs. The breakpoints of these LOH events were different in the three strains: KO_5_6_E (about 135 kb); KO_5_9_H (about 127 kb); KO_75_2_XX_H (about 36 kb). According to Wei *et al.* (2007), in YJM789, an approximately 30 kb segment derived from the left end of chromosome VI is translocated to the left end of chromosome X. From these data, the W303-1A SNPs are at the right end of chromosome VI, whereas the YJM789 SNPs are at the right end of chromosome X. Since sub-telomeric repeats are difficult to assemble, it is also possible that YJM789 has two copies of the 30 kb segment, one on VI and one on X. An LOH event on the left arm of chromosome X in which YJM789-derived sequences are lost will result in an apparent deletion of YJM789-specific sequences near the left telomere of VI. It should also be noted that there are only eight SNPs on the microarray from the 30 kb segment.

KO_75_2_XX_I (*rnh201*Δ); KO_5_9_I (*rnh1*Δ *rnh201*Δ). These strains both contain apparent terminal duplications of YJM789-derived sequences on the left arm of

chromosome XV with the starting point of the duplication near coordinate 22 kb. Both strains have terminal LOH events on the left end of chromosome IX, resulting in duplication of YJM789-derived SNPs and loss of W303-1A-derived SNPs. The breakpoints for the LOH event are near coordinate 105 kb for KO_75_2_XX_I and near coordinate 336 for KO_5_9_I. The sub-telomeric regions of chromosomes IX and XV share considerable homology in both the YJM789 and W303-1A/S288c genomes. In the S288c genomes, sequences from chromosome XV with coordinates about 22-31 kb share extensive homology with sequences from chromosome IX located between coordinates 17-26 kb. However, according the SGD database and our analysis, the oligonucleotides that have an elevated signal on chromosome XV are not in the region of the genome that is repeated on chromosome IX in either the W303-1A or the YJM789 genomes. For example, in isolate KO_75_2_XX_I, the YJM789-specific oligonucleotide at position 22005 clearly has an elevated level of hybridization. The sequence of this oligonucleotide, however, is not present on chromosome IX. Similarly, in isolate KO_5_9_I, the level of hybridization to the YJM789-specific oligonucleotide 19925 is elevated, although the sequence of this oligonucleotide is not on chromosome IX.

Although we do not have a definitive explanation of these observations, one possibility is that the YJM789 isolate used in our studies has a derivative of chromosome IX in which the terminal 22 kb of chromosome X replaces the terminal 16 kb of chromosome IX. Such a derivative could be formed as a consequence of a break-induced replication event in which a broken end of the YJM789-derived copy of chromosome IX duplicates a portion of the YJM789-derived chromosome XV homolog. The initiation point of this invasion would be in the region of shared homology. In diploid strains with this derivative, an LOH event on chromosome IX, occurring centromere-proximal to the duplicated region, would duplicate both YJM789-related SNPs on chromosome XV and cause a duplication of YJM789-derived sequences from the

terminal region of chromosome XV. It should be noted that an LOH event on chromosome XV that results in LOH for the terminal repeated sequences would not be annotated as an LOH event on chromosome IX since the first oligonucleotide used to diagnose LOH for chromosome IX is located at coordinate 25 kb. Although this model is consistent with our observations, our observations could also reflect an assembly error of the genomic sequences.

KO_5_6_K (*rnh1*Δ *rnh201*Δ). In this isolate, there is an apparent duplication of YJM789-derived sequences on the left arm of chromosome XVI with a transition point between coordinates 23222 and 26225. There were also terminal LOH events on several chromosome arms including the right arm of chromosome XIII, duplicating YJM789-derived SNPs; the LOH event on XV has a transition between coordinates 889 and 892 kb. The left arm of chromosome XVI (25.8-26.4 kb) shares homology with the right arm of chromosome XIII (coordinates 917.5 kb to 917.8 kb), although the oligonucleotides that have increased levels of hybridization on chromosome XVI in KO_5_6_K are not annotated as duplicated on any other homolog. One explanation of the data is that a break within the shared homology occurred on chromosome XIII that was repaired by a BIR event involving chromosome XVI. A subsequent LOH event on XIII could result in the observed apparent duplication of XVI sequences as well as the terminal LOH event duplicating YJM789-derived SNPs. Alternatively, there may be an incorrect assembly or annotation of the sub-telomeric sequences in the databases.

KO_5_9_D (*rnh1*Δ *rnh201*Δ). In this isolate, there is an apparent deletion of W303-1A-derived sequences on the right arm of chromosome I with a transition point between coordinates 195120 and 203572. There is also a terminal LOH event on the right arm of chromosome VIII, resulting in loss of W303-1A-derived sequences, with a transition point between 199775 and 207066. In the S288c genome, there is a large region of conserved

homology that includes the coordinates 207-227 kb on the right end of chromosome I and coordinates 528-556 kb on the right end of VIII. Several of the SNPs located near the right end of chromosome I (for example, oligonucleotide 208214; Table S3, St. Charles *et al.*, 2012) are duplicated on the right end of VIII. Therefore, an LOH event that causes loss of W303-1A-derived SNPs and duplication of YJM789-derived SNPs will result in an apparent deletion of W303-1A sequences from the right end of chromosome I.

Supplemental Figure Legends

Figure S1. Patterns of LOH in sub-cultured strains. Each line represents markers in a diploid isolate. Green indicates heterozygous SNPs; red, homozygous W303-1A-derived SNPs; black, homozygous YJM789-derived SNPs. The yellow circle shows the centromere. Each transition between heterozygous and homozygous SNPs or between two regions with different homozygous SNPs is labeled with a lower case letter. Classes a1-a4 are simple terminal LOH events. In Classes a6-a8, the two transitions (one marked with an asterisk) are separated by distances that are two standard deviations longer than the median length of a mitotic conversion tract. The two transitions are, therefore, likely to reflect two different recombination events. Classes b1 and b2 represent simple interstitial LOH events (gene conversions), whereas in Classes b3-b5, the conversion event is interrupted by a region of heterozygosity. Classes f1-f12 represent terminal LOH events with complex patterns of associated LOH events. Only Classes a1-a4, b1, and b2 were used for our association studies.

Figure S2. Deletions and duplications in sub-cultured strains. This diagram shows the patterns of deletions and duplications in diploid isolates. As in Fig. S1, the green line indicates heterozygous SNPs, and yellow circles show the centromere. The deletion or duplication is shown as a line that is half as wide as the green lines. The Classes dd9 and dd10 show interstitial duplications of W303-1A-derived and YJM789-derived SNPs, respectively. The Class dd12 shows an interstitial deletion in which W303-1A-derived SNPs were removed. The coordinates for the transitions are in Table S6.

Figure S3. Aneuploidy events in sub-cultured strains. Trisomic, but not monosomic, aneuploid events were observed in our studies. For each chromosome, we indicate whether the strain has W303-1A-derived SNPs (red) or YJM789-derived SNPs (black). Note that many of these aneuploid events are associated with recombination on one or more chromosomes.

Figure S4. Locations of LOH events in the sub-cultured *rnh201Δ* strain. Each of the sixteen chromosomes is shown as a thin black horizontal line with SNPs shown as very short vertical yellow lines. The centromeres are represented by black ovals. Red and blue bars show regions of interstitial LOH in which the W303-1A-derived SNPs became homozygous and the YJM789-derived SNPs became homozygous, respectively. Black arrows indicate the positions of terminal LOH events that were unassociated with a gene conversion event. Red arrows and blue arrows indicate terminal LOH events that were associated with a conversion that made W303-1A-derived SNPs and YJM789-derived SNPs homozygous, respectively. Triangles indicate deletions (red for a deletion of W303-1A-derived sequences and blue for a deletion of YJM789-derived sequences) and inverted triangles indicate duplications (same color code as for deletions).

Figure S5. Location of LOH events in the sub-cultured *rnh1Δ rnh201Δ* strain. The mapped events are shown with the same code as in Fig. S4.

Figure S6. Location of LOH events in the sub-cultured *rnh201Δ pol2-M644L* strain. The mapped events are shown with the same code as in Fig. S4.

Figure S7. Patterns of LOH in sectored colonies. In this depiction, each sectored colony is represented by a pair of lines with the red sector shown as the top line. We use the same color code for heterozygous and homozygous regions as in Fig. S1. Other features of these patterns are described in the main text.

Supplemental Tables.

Table S1. Strain list.

Table S2. Plasmid list.

Table S3. Primers list.

Table S4. Strains used in different assays of LOH.

Table S5. LOH events in sub-cultured strains.

Table S6. Deletion/duplication events in sub-cultured strains.

Table S7. Trisomy events in sub-cultured strains.

Table S8. LOH events on chromosome IV in sectored colonies.

Table S9. References used to determine the locations of genomic elements.

Table S10. Number of genomic elements represented on the microarrays.

Table S11. Association of LOH events in sub-cultured strains with various genomic elements.

Table S12. Association of LOH events in sectored colonies with various genomic elements on the right arm of chromosome IV.

References

Altman, D. G. (1990) *Practical Statistics for Medical Research*. (Chapman and Hall/CRC Texts in Statistical Science).

Aksenova, A.Y, P.W. Greenwell, M. Dominska, A.A. Shishkin, J.C. Kim, T.D. Petes, and S.M. Mirkin (2013) Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proc Natl Acad Sci USA* 110(49): 19866-19871.

Azvolinsky, A., P.G. Giresi, J.D. Lieb, and V.A. Zakian (2009) Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. Mol Cell 34: 722-34.

Benjamini Y. and Y. Hochberg (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc A Sta* 57(1): 289-300.

Capra, J.A., K. Paeschke, M. Singh, V.A. Zakian (2010) G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in*Saccharomyces cerevisiae*. *PLoS Comput Biol* 6: e1000861

Chan, Y. A., M.J. Aristizabal, Y.T. Lu Phoebe, Z. Luo, A. Hamza, M.S. Kobor, P.C. Stirling, P. Hieter (2014) Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-Chip. *PLoS Genet* 10(4): e1004288.

Engel S.R., F.S. Dietrich, D.G. Fisk, G. Binkley, R. Balakrishnan, M.C. Costanzo, S.S. Dwight, B.C. Hitz, K. Karra, R.S. Nash, S. Weng, E.D. Wong, P. Lloyd, M.S. Skrzypek, S.R. Miyasato, M. Simison, J.M. Cherry (2013) The Reference Genome Sequence of Saccharomyces cerevisiae: Then and Now. *G3* pii: g3.113.008995v1. doi: 10.1534/g3.113.008995.

Fachinetti, D., R. Bermejo, A. Cocito, S. Minardi, Y. Katou, Y. Kanoh, K. Shirahige, A. Azvolinsky, V.A. Zakian, M. Foiani (2010) Replication termination at eukaryotic

chromosomes is mediated by Top2 and occurs at genomic loci containing pause elements. *Molecular Cell*. 39(4): 595-605.

Goldstein, A.L., and J.H. McCusker (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* 15(14): 1541-1553.

Gueldener, U., J. Heinisch, G.J. Koehler, D. Voss, and J.H. Hegemann (2002) A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 30(6): e23.

Hershman, S.G., Q. Chen, J.Y. Lee, M.L. Kozak, P. Yue, L.S. Wang, and F.B. Johnson (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 36(1): 144-156.

Hobson, D. J., W. Wei, L.M. Steinmetz, J.Q. Svejstrup (2012) RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48(3): 365-374.

Kim, N., J. Cho, Y.C. Li, and S. Jinks-Robertson (2013) RNA:DNA hybrids initiate quasi-palindrome associated mutations in highly transcribed yeast DNA. *PLoS Genet* 9(11): e1003924.

Lea, D.E. and C.A. Coulson (1949) The distribution of number of mutants in a bacterial population. *J Genet* 49: 264-285.

Lee, P.S., P.W. Greenwell, M. Dominska, M. Gawel, M. Hamilton, and T.D. Petes (2009) A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. *PLoS Genet* 5(3): e1000410.

Lisnic, B., I.K. Svetec, H. Saric, I. Nikolic, and Z. Zgaga (2005) Palindrome content of the yeast *Saccharomyces cerevisiae* genome. *Curr Genet* 47: 289-297.

Longtine, M.S., A. McKenzie III, D.J. Demarini, N.G. Shah, A. Wach, A. Brachat, P. Philippsen, and J.R. Pringle (1998) Additional modules for versatile and economical

PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14: 953-961.

Mortimer, R.K. and J.R. Johnston (1986) Genealogy of principal strains of the yeast genetic stock center. *Genetics* 113(1): 35-43.

Nick McElhinny, S.A., D. Kumar, A.B. Clark, D.L. Watt, B.E. Watts, E.B. Lundström, E. Johansson, A. Chabes, and T.A. Kunkel (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6(10): 774-781.

Song, W., M. Dominska, P.W. Greenwell, and T.D. Petes (2014) Genome-wide high-resolution mapping of chromosome fragile sites in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 111(21): e2210-e2218.

St. Charles, J., E. Hazkani-Covo, Y. Yin, S.L. Andersen, F.S. Dietrich, P.W. Greenwell, E. Malc, P. Mieczkowski, and T.D. Petes (2012) High-resolution genome-wide analysis of irradiated (UV and γ-rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events. *Genetics* 190(4): 1267-1284.

St. Charles, J. and T.D. Petes (2013) High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1Mb arm of yeast chromosome IV. *PLoS Genet* 9(4): e1003434.

Thomas, B.J. and R. Rothstein (1989) Elevated recombination rates in transcriptionally active DNA. *Cell* 56(4): 619-630.

Vernon, M., K. Lobachev, and T.D. Petes (2008) High rates of “unselected” aneuploidy and chromosome rearrangements in *tel1 mec1* haploid yeast strains. *Genetics* 179(1): 237-247.

Wei, W., J.H. McCusker, R.W. Hyman, T. Jones, Y. Ning, Z. Cao, Z. Gu, D. Bruno, M. Miranda, M. Nguyen, J. Wilhelmy, C. Komp, R. Tamse, X. Wang, P. Jia, P. Luedi, P.J. Oefner, L. David, F.S. Dietrich, Y. Li, R.W. Davis, and L.M. Steinmetz (2007)

Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc Natl Acad Sci USA* 104(31): 12825-12830.

Zhao, X., E.G. Muller, and R. Rothstein (1998) A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Mol Cell* 2: 329-340.

Table S1. Strain list

Strain Name	Relevant genotype	Construction or source	Genotype	Strain background
S288c	Wild type	Mortimer and Johnston 1986	*MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6*	S288c
W1588-4c	Wild type	Zhao *et al.* 1998	*MAT**a** leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5*	W303-1A
W303-1A	Wild type	Thomas and Rothstein 1989	*MAT**a** leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5*	W303-1A
SMY710	Wild type	Aksenova *et al.* 2013	*MAT**a** leu2-Δ1 trp1-Δ63 ura3-52 his3-200 ade2Δ::kanMX*	S288c
MV70	Wild type	Vernon *et al.* 2008	*MAT**a**/MATα trp1-1/trp1-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 ura3-1/ura3-1 ade2-1/ade2-1 can1-100/CAN1 hom3-10/HOM3 rad5/RAD5 tel1Δ::kanMX/TEL1 mec1-21/MEC1*	W303-1A/ W303-1A
PG308	Wild type	MV70 transformed with PCR fragment amplified from pAG25 using primers Tel1NatF and Tel1NatR.	*MAT**a**/MATα trp1-1/trp1-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 ura3-1/ura3-1 ade2-1/ade2-1 can1-100/CAN1 hom3-10/HOM3 rad5/RAD5 tel1Δ::natMX/TEL1 mec1-21/MEC1*	W303-1A/W303-1A
PG309(2)	Wild type	PG308 transformed with PCR fragment amplified from pFA6-kanMXpGAL(x3HA) using primers MRC1pgalF and MRC1pgalR	*MAT**a**/MATα trp1-1/trp1-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 ura3-1/ura3-1 ade2-1/ade2-1 can1-100/CAN1 hom3-10/HOM3 rad5/RAD5 tel1Δ::nat/TEL1 mec1-21/MEC1 MRC1/pGAL-MRC1-kanMX*	W303-1A/ W303-1A
PG309(2)-4a	*tel1*Δ	Spore from PG309(2)	*MATα RAD5 leu2-3,112 his3-11,15 trp1-1 ade2-1 tel1Δ::natMX*	W303-1A
JSC19-1	Wild type	St. Charles and Petes 2013	*MATα ade2-1 ura3 gal2 ho::hisG can1Δ::natMX*	YJM789
JSC21-1	Wild type	St. Charles and Petes 2013	*MATα ura3 gal2 ho::hisG ade2-1 can1Δ::natMX IV1510386::SUP4-o*	YJM789
JSC12	Wild type	St. Charles and Petes 2013	*MAT**a** leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1Δ::natMX RAD5 IV1510386::kanMX-can1-100*	W303-1A
SJR3585	Wild type	Spore from diploid PG309(2)-4a x JSC12	*MAT**a** RAD5 leu2-3,112 trp1-1 his3-11,15 ura3-1 ade2-1 IV1510386::kanMX-can1-100*	W303-1A
SJR3615-4	*rnh201*Δ	SJR3585 transformed with PCR fragment amplified from pSR955 using primers	*MAT**a** leu2-3,112 his3-11,15 trp1-1 ura3-1 ade2-1 RAD5 rnh201Δ::loxP-hphMX-loxP IV1510386::kanMX-can1-100*	W303-1A

		RNH2kanF and RNH2kanR		
SJR3625-9B	*rnh1*Δ *rnh201*Δ	SJR3615-4 transformed with PCR fragment amplified from pSR879 using primers RNH1kanF and RNH1kanR	*MAT***a** *leu2-3,112 his3-11,15 trp1-1 ura3-1 ade2-1 RAD5 rnh1Δ::loxP-natMX-loxP rnh201Δ::loxP-hphMX-loxP IV1510386::kanMX-can1-100*	W303-1A
SJR3616-3	*rnh201*Δ	SJR3586 transformed with PCR fragment amplified from pSR955 using primers RNH2kanF and RNH2kanR.	*MATα ura3 gal2 ho::hisG ade2-1 can1Δ::natMX rnh201Δ::loxP-hphMX-loxP IV1510386::SUP4-o*	YJM789
SJR3626-3	*rnh1*Δ *rnh201*Δ	SJR3616-3 transformed with PCR fragment amplified from pUG6 using RNH1kanF and RNH1kanR	*MATα ura3 gal2 ho::hisG ade2-1 can1Δ::natMX rnh201Δ::loxP-hph-loxP rnh1Δ::loxP-kanMX-loxP IV1510386::SUP4-o*	YJM789
YJM799	Wild type	John McCusker	*MATα ura3 gal2 ho::hisG*	YJM789
YJM790	Wild type	John McCusker	*MAT***a** *ho::hisG lys2 gal2*	YJM789
KOK3	Wild type	W1588-4c transformed with PCR fragment amplified from pUG6 using primers Forward Sequence and Reverse Sequence.	*MAT***a** *can1-100 trp1-1 ade2-1 his3-11,15 leu2-3,112 ura3Δ::loxP-kanMX-loxP*	W303-1A
KO5	*rnh1*Δ *rnh201*Δ	SJR3625-9B x SJR3626-3	*MAT***a***/MATα ho::hisG/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP rnh1Δ::loxP-natMX-loxP/rnh1Δ::loxP-kanMX-loxP IV1510386::kanMX-can1-100/IV1510386::SUP4-o*	W303-1A/YJM789
KO30	Wild type	SJR3626-3 x YJM790	*MAT***a***/MATα ho::hisG/ho::hisG gal2/gal2 ade2-1/ADE2 URA3/ura3 lys2/LYS2 CAN1/can1Δ::natMX RNH201/rnh201Δ::loxP-hphMX-loxP RNH1/rnh1Δ::loxP-kanMX-loxP IV1510386/IV1510386::SUP4-o*	YJM789
KO32	*rnh1*Δ *rnh201*Δ	Spore from KO30	*MAT***a** *can1Δ::natMX ade2-1 ura3 ho::hisG gal2 rnh1::loxP-kanMX-loxP rnh201::loxP-hphMX-loxP*	YJM789
KO35	Wild type	JSC19-1 transformed with PCR fragment amplified from S288c using primers IVURA3F and IVURA3R.	*MATα ade2-1 ura3 gal2 ho::hisG can1Δ::nat IV1510386::URA3*	YJM789

KO36	Wild type	Cre-expressing plasmid pSH47 used to excise the kanMX marker in KOK3	*MAT**a** can1-100 trp1-1 ade2-1 his3-11,15 leu2-3,112 ura3Δ::loxP*	W303-1A
KO49	Wild type	KO32 x KO35	*MAT**a**/MATα ade2-1/ade2-1 ura3/ura3 gal2/gal2 ho::hisG/ho::hisG can1Δ::nat/can1Δ::nat rnh1Δ ::loxP-kanMX-loxP/RNH1 rnh201Δ::loxP-hphMX-loxP/RNH201 IV1510386/IV1510386::URA3*	YJM789/ YJM789
KO52	Wild type	SJR3625-9B x W303-1B	*MAT**a**/MATα leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 trp1-1/trp1-1 ura3-1/ura3-1 ade2-1/ade2-1 CAN1/can1-100 IV1510386::kanMX-can1-100/IV1510386 rnh1Δ::loxP-natMX-loxP/RNH1 rnh201Δ::loxP-hphMX-loxP/RNH201*	W303-1A/ W303-1A
KO57	*rnh1*Δ	Spore of KO49	*MATα ade2-1 ura3 gal2 ho::hisG can1Δ::nat IV1510386::URA3 rnh1Δ::loxP-kanMX-loxP*	YJM789
KO63	*rnh1*Δ *rnh201*Δ	Spore from KO52	*MATα leu2-3,112 his3-11,15 trp1-1 ura3-1 ade2-1 can1-100 rnh1Δ::loxP-natMX-loxP rnh201Δ::loxP-hphMX-loxP*	W303-1A
KO70	*rnh1*Δ	Spore of KO52	*MAT**a** leu2-3,112 his3-11,15 trp1-1 ura3-1 ade2-1 can1-100 rnh1Δ::loxP-nat-loxP*	W303-1A
KO73	*rnh1*Δ	KO70 x KO57	*MAT**a**/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 rnh1Δ::loxP-natMX-loxP/rnh1Δ::loxP-kanMX-loxP IV1510386::URA3/IV1510386*	W303-1A/ YJM789
KO75	*rnh201*Δ	SJR3616-3 x SJR3615-4	*MAT**a**/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP IV1510386::URA3/IV1510386*	W303-1A/ YJM789
KO119	Wild type	YJM799 transformed with transformed with PCR fragment amplified from SMY710 using primers ADE2_XV_F and ADE2_XV_R	*MATα ura3 gal2 ho::hisG ade2Δ::kanMX*	YJM789
KO124	Wild type	KO119 transformed with transformed with PCR fragment amplified from	*MATα ade2Δ::kanMX ura3 ho::hisG gal2 IV1495420::ADE2*	YJM789

		S288c using primers IVADE2_3_F and IVADE2_3_R		
KO125	Wild type	KO124 x KO32	*MAT**a**/MATα can1Δ::natMX/CAN1 ade2-1/ade2Δ::kanMX ura3/ura3 ho::hisG/ho::hisG gal2/gal2 rnh1Δ::loxP-kanMX-loxP/RNH1 rnh201Δ::loxP-hphMX-loxP/RNH201 IV1495420/IV1495420::ADE2*	YJM789/ YJM789
KO127	*rnh1*Δ *rnh201*Δ	Spore from KO125	*MATα ade2-1 ura3 ho::hisG gal2 can1Δ::natMX rnh1Δ::loxP-kanMX-loxP rnh201Δ::loxP-hphMX-loxP IV1495420::ADE2*	YJM789
KO128	*rnh1*Δ	Spore from KO125	*MATα ade2-1 ura3 ho::hisG gal2 can1Δ::natMX rnh1Δ::loxP-kanMX-loxP IV1495420::ADE2*	YJM789
KO131	*rnh201*Δ	Spore from KO125	*MATα ade2-1 ura3 ho::hisG gal2 rnh201Δ::loxP-hphMX-loxP IV1495420::ADE2*	YJM789
KO132	*rnh1*Δ *rnh201*Δ	KO127 x SJR3625-9B	*MAT**a**/MATα leu2-3,112/LEU2 his3-11,15/HIS3 trp1-1/TRP1 ura3-1/ura3 ade2-1/ade2-1 rnh1Δ::loxP-natMX-loxP /rnh1Δ::loxP-kanMX-loxP rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP ho::hisG/ho::hisG GAL2/gal2 IV1510386::kanMX-can1-100/IV1510386 IV1495420/IV1495420::ADE2*	W303-1A/ YJM789
KO135	*rnh201*Δ	KO130 x SJR3615-4	*MAT**a**/MATα leu2-3,112/LEU2 his3-11,15/HIS3 trp1-1/TRP1 ura3-1/ura3 HO/ho::hisG ade2-1/ade2-1 GAL2/gal2 CAN1/can1Δ::natMX IV1510386::kanMX-can1-100/IV1510386 rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP IV1495420/IV1495420::ADE2*	W303-1A/ YJM789
KO171	Wild type	KOK3 transformed with transformed with PCR fragment amplified from pUG72 using primers KL_URA3_F and KL_URA3_R.	*MAT**a** can1-100 trp1-1 ade2-1 his3-11,15 leu2-3,112 ura3Δ::loxP-kanMX-loxP IV1495420::loxP-URA3Kl-loxP*	W303-1A
KO172	Wild type	KO171 x KO63	*MAT**a**/MATα can1-100/can1-100 trp1-1/trp1-1 ade2-1/ade2-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 ura3Δ::loxP-kanMX-loxP/ura3-1 RNH201/rnh201Δ::loxP-hphMX-loxP RNH1/rnh1Δ::loxP-natMX-loxP IV1495420::loxP-URA3Kl-loxP/IV1495420*	W303-1A/ W303-1A

KO175	*rnh1*Δ *rnh201*Δ	Spore from KO172	*MAT***a** *can1-100 trp1-1 ade2-1 ura3-1 leu2-3,112 his3-11,15 rnh201Δ::loxP-hph-loxP rnh1Δ::loxP-natMX-loxP IV1495420::loxP-URA3KI-loxP*	W303-1A
KO176	*rnh1*Δ *rnh201*Δ	Spore from KO172	*MATα can1-100 trp1-1 ade2-1 leu2-3,112 his3-11,15 ura3-1 rnh201Δ::loxP-hphMX-loxP rnh1Δ::loxP-natMX-loxP IV1495420::loxP-URA3KI-loxP*	W303-1A
KO179	*rnh201*Δ	Spore from KO172	*MAT***a** *can1-100 trp1-1 ade2-1 ura3-1 leu2-3,112 his3-11,15 rnh201Δ::loxP-hph-loxP IV1495420::loxP-URA3KI-loxP*	W303-1A
KO185	*rnh1*Δ	Spore from KO172	*MAT***a** *can1-100 trp1-1 ade2-1 ura3-1 leu2-3,112 his3-11,15 rnh1Δ::loxP-natMX-loxP IV1495420::loxP-URA3KI-loxP*	W303-1A
KO187	*rnh1*Δ	KO185 x KO128	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/can1Δ::natMX rnh1Δ::loxP-natMX-loxP/rnh1Δ::loxP-kanMX-loxP IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789
KO188	*rnh201*Δ	KO179 x KO131	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/can1Δ::natMX rnh201Δ::loxP-natMX-loxP/rnh201Δ::loxP-kanMX-loxP IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789
KO189	*rnh1*Δ *rnh201*Δ	KO175 x KO127	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP rnh1Δ::loxP-natMX-loxP/rnh1Δ::loxP-kanMX-loxP IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789
KO198	Wild type	KO171 x KO124	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2Δ::kanMX ura3Δ ::loxP-kanMX-loxP/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789
KO200	*pol2-M644L*	*pol2-M644L* introduced into KO36 by two-step allele replacement using AgeI-digested p173-pol2-M644L	*MAT***a** *can1-100 trp1-1 ade2-1 his3-11,15 leu2-3,112 ura3Δ::loxP pol2-M644L*	W303-1A
KO201	Wild type	KO200 x KO176	*MAT***a***/MATα can1-100/can1-100 trp1-1/trp1-1 ade2-1/ade2-1 his3-11,15/his3-11,15 leu2-3,112/leu2-3,112*	W303-1A/ W303-1A

			ura3Δ::loxP/ura3-1 pol2-M644L/POL2 RNH1/rnh1Δ::loxP-natMX-loxP rnh201Δ::loxP-hphMX-loxP/RNH201 IV1495420/IV1495420::ADE2	
KO204	*rnh201*Δ *pol2-M644L*	Spore from KO201	*MAT***a** *can1-100 trp1-1 ade2-1 his3-11,15 leu2-3,112 ura3-1 pol2-M644L rnh201Δ::loxP-hphMX-loxP IV1495420::loxP-URA3K.L-loxP*	W303-1A
KO207	*rnh201*Δ *pol2-M644L*	*pol2-M644L* introduced into KO131 by two-step allele replacement using *Age*I-digested p173-pol2-M644L	*MATα ade2-1 can1Δ::natMX ura3 ho::hisG gal2 pol2-M644L rnh201Δ::loxP-hphMX-loxP IV1495420::ADE2*	YJM789
KO213	*pol2-M644L*	KO124 after two-step transplacement using *Age*I-digested p173-pol2-M644L	*MATα ade2Δ::kanMX ura3 ho::hisG gal2 IV1495420::ADE2 pol2-M644L*	YJM789
KO214	Wild type	KO207 x KO32	*MAT***a***/MATα can1Δ::natMX/CAN1 ade2-1/ade2-1 ura3/ura3 ho::hisG/ho::hisG gal2/gal2 rnh1Δ::loxP-kanMX-loxP/RNH1 rnh201Δ::loxP-hphMX-loxP/RNH201 POL2/pol2M644L IV1495420/IV1495420::ADE2*	YJM789/ YJM789
KO218	*rnh201*Δ *pol2-M644L*	Spore from KO214	*MATα ade2-1 ura3 gal2 ho::hisG pol2-M644L can1::ΔnatMX rnh201Δ::loxP-hphMX-loxP IV1495420::ADE2*	YJM789
KO234	*pol2-M644L/ pol2-M644L*	KO213 x KO200	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/CAN1 pol2-M644L/pol2-M644L IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789
KO244	*rnh201*Δ *pol2-M644L*	KO218 x KO204	*MAT***a***/MATα HO/ho::hisG ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 trp1-1/TRP1 his3-11,15/HIS3 leu2-3,112/LEU2 can1-100/can1Δ::natMX rnh201Δ::loxP-hphMX-loxP/rnh201Δ::loxP-hphMX-loxP pol2-M644L/pol2-M644L IV1495420::loxP-URA3KI-loxP/IV1495420::ADE2*	W303-1A/ YJM789

Table S2. Plasmid list

Plasmid	Relevant feature	Source	Strains constructed
pFA6-kanMXpGAL(x3HA)	*kanMX-pGAL* cassette	Longtine *et al.* 1998	PG309(2)
p173-pol2-M644L	*pol2-M644L* allele	Nick McElhinny *et al.* 2010	KO200, KO207, KO213
pAG25	*natMX*	Goldstein and McCusker 1999	PG308
pSH47	*pGAL1-Cre*	Gueldener *et al.* 2002	KO36
pUG72	*loxP-URA3Kl-loxP*	Gueldener *et al.* 2002	KO171
pSR879	*loxP-natMX-loxP*	Kim *et al.* 2013	SJR3625-9B
pSR955	*loxP-hphMX-loxP*	Kim *et al.* 2013	SJR3615-4, SJR3616-3
pUG6	*loxP-kanMX-loxP*	Gueldener *et al.* 2002	SJR3626-3, KOK3

Table S3. Primers list

Primer	Sequence (5′ to 3′)	Strains constructed (purpose)
RNH1KANF	ATGGCAAGGCAAGGGAACTTCTACGCGGTTAGAAAGGGCAGGGAAA CTGGGATCTATAATCAGCTGAAGCTTCGTACG	SJR2626-3, SJR3625-9B (construct *RNH1* deletion)
RNH1KANR	GCATTATCGTCTAGATGCTCCTTTCTTCGCCAGAAAATCTGCCATTTC ATTTCCTGGATCAGGCCACTAGTGGATCTG	SJR2626-3, SJR3625-9B (construct *RNH1* deletion)
RNH1UPF	TGGCAGCACAATAATACACG	KO127, KO128, SJR3626-3, SJR3625-9b (confirm *rnh1*Δ)
RNH1DWNR	CACGCTTATAGATAGTTATCG	KO127, KO128, SJR3626-3, SJR3625-9b (confirm *rnh1*Δ)
RNH2KANF	CTAATGAGAGTGTCGAAAAACCTTGAAAACAACTACTGCACACCAAAT TGATACGATTAACAGCTGAAGCTTCGTACG	SJR3615-4, SJR3616-3 (*RNH201* deletion)
RNH2KANR	GCTTCACGGATAGTAGAAACGGCAAAGCATAGTAGCAGATGACTTGT ATGAGTTATTGAAAGGCCACTAGTGGATCTG	SJR3615-4, SJR3616-3 (*RNH201* deletion)
RNH2UPF	TTGCGACGCCTGCCAATGC	SJR3615-4, SJR3616-3 (confirm *rnh201*Δ)
RNH2DWR	TCGTTCCGGTTGGTTGTCTC	SJR3615-4, SJR3616-3 (confirm *rnh201*Δ)
KL_URA3_F	CGGGTAGAATCAATGCAATCAGTGGTAATTATCTAGATGACGTCCTTT ATGACCTTGACACCCCTGCAGCTGAAGCTTCGTACG	KO171 (insertion of *URA3KI* at end of chr. IV)
KL_URA3_R	TGTGGTGACAACCTAACCCTTTCGTTGATACTAGTTTGAAGGTTATCA ATATCCTGAATTAGAGTTGTGGAGGCCACTAGTGGATCTG	KO171 (insertion of *URA3KI* at end of chr. IV)
IVADE2_3_F	CGGGTAGAATCAATGCAATCAGTGGTAATTATCTAGATGACGTCCTTT ATGACCTTGACACCCCTGGTTGAGAAGCCGAGAATTTTGTA	KO124 (insertion of *ADE2* at end of chr. IV)
IVADE2_3_R	TGTGGTGACAACCTAACCCTTTCGTTGATACTAGTTTGAAGGTTATCA ATATCCTGAATTAGAGTTGTGGTCCTCGGTTCTGCATTGAG	KO124 (insertion of *ADE2* at end of chr. IV)
IVURA3F	GCTTTACAGGACCTATTTTTCATACGTTATGCACTTCATTCTTTTTGTC GGTTTGATACCAGCAGAATCTAACGCTAGAGCAGACGCTCAT	KO35 (insertion of *URA3* at end of chr. IV)
IVURA3R	AAGCGCTGCTGCGTTTTCGAGGTATGGCTTCTGCCGGGCTAACGTTC AAATTAAAGGAACAGATTCCCGGGTAATAACTGA	KO35 (insertion of *URA3* at end of chr. IV)
EXT1510386F	CATTGGAGCGAATGATGACG	KO35 (confirmation of *URA3* insertion on chr. IV)
EXT1510386R	TGTGCAATCGTTGTCAAGTTGG	KO35 (confirmation of *URA3* insertion on chr. IV)
Forward Sequence	TGCCCAGTATTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAA ACGAAGATAAATCCAGCTGAAGCTTCGTACG	KOK3 (deletion of *URA3* locus)
Reverse	TTAAATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTACTTAT	KOK3 (deletion of *URA3* locus)

Sequence	AATACAGTTTTAGGCCACTAGTGGATCTG	
ADE2_XV_R	GGTGCGTAAAATCGTTGGAT	KO119, KO127, KO128 (to insert *ade2Δ::kanMX* in KO119 and to determine whether KO127 and KO128 had *ade2-1* or *ade2Δ::kanMX*)
ADE2_XV_F	ATCCTCGGTTCTGCATTGAG	KO119, KO127, KO128 (to insert *ade2Δ::kanMX* in KO119 and to determine whether KO127 and KO128 had *ade2-1* or *ade2Δ::kanMX*)
Pol2DigestF1	GAAAAGCCACAGCACCTTTC	KO200, KO207, spores of KO214 and KO201 (confirmation of *pol2-M644L)*
Pol2DigestR1	TTGGACAGATTTCCCTTCCA	KO200, KO207, spores of KO214 and KO201 (confirmation of *pol2-M644L)*
MATaF	ACTCCACTTCAAGTAAGAGTTTG	Many strains (diagnosis of mating type)
MATalphaF	GCACGGAATATGGGACTACTTCG	Many strains (diagnosis of mating type)
MATR	AGTCACATCAAGATCGTTTATGG	Many strains (diagnosis of mating type)
extF3	AATGCGGGTAGAATCAATGC	KO124, KO171 (confirmation of *ADE2* or *URA3KI* insertion on chr. IV)
extR3	AGGTGATGGGAAATCGAGTG	KO124, KO171 (confirmation of *ADE2* or *URA3KI* insertion on chr. IV)
KANF222	AATTTATGCCTCTTCCGACC	KOK3 (confirmation of *URA3* deletion)
URAR1128	GAAATCATTACGACCG	KOK3 (confirmation of *URA3* deletion)
Tel1NATF	ATTCGAAAAAAAAGCCTTCAAAGAAAAGGGAAATCAGTGTAACATAGACGCGTACGCTGCAGGTCGAC	PG308 (replacement of *tel1::kanMX* with *tel1::natMX*)
Tel1NATR	TTCGTATTTCTATAAACAAAAAAAAGAAGTATAAAGCATCTGCATAGCAAATCGATGAATTCGAGCTCG	PG308 (replacement of *tel1::kanMX* with *tel1::natMX*)
MRC1pgalF	GGAAGTTCGTTATTCGCTTTTGAACTTATCACCAAATATTGAATTCGAGCTCGTTTAAAC	PG309(2) (to insert the *GAL1* promoter and *kanMX* upstream of *MRC1*)
MRC1pgalR	TTGCAGTCAACGAGGACAAAGCATGCAAGGCATCATCCATGCACTGAGCAGCGTAATCTG	PG309(2) (to insert the *GAL1* promoter and *kanMX* upstream of *MRC1*)

Table S4. Strains used in different assays of LOH

Genotype	Sub-culturing	Sectoring assay	RCO mapping in sectored colonies	Rate of 5FOA resistance
wild-type	KO198	KO198	KO198	KO198
*rnh1*Δ	KO73	KO187	KO187	KO187
pol2-M644L	KO234	KO234	KO234	KO234
*rnh201*Δ	KO75	KO188	KO188 and KO135	KO188
*rnh201*Δ *pol2-M644L*	KO244	KO244	KO244	KO244
*rnh1*Δ *rnh201*Δ	KO5	KO189	KO132	KO189

Tables S5-S8

Available for download as Excel files at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.182725/-/DC1

Table S5 LOH events in sub-cultured strains

Table S6 Deletions-Duplications, sub-cultured strains

Table S7 Trisomy

Table S8 LOH events on chromosome IV in sectored colonies

Table S9. References used to determine locations of genomic elements.[1]

Genomic elements	Data source
Ty element	SGD; YeastMine
Solo LTR	SGD; YeastMine
Centromeres	SGD; YeastMine
Intron-containing genes	SGD; YeastMine
ARS elements	SGD; YeastMine
tRNA genes	SGD; YeastMine
Long genes	SGD; YeastMine
Regions of high transcription	SGD; YeastMine
Regions of low transcription	SGD; YeastMine
ORFs with high GC content	SGD; YeastMine
High G content on the non-transcribed strand	SGD; YeastMine
Sites of Rbp3 accumulation in S phase	Fachinetti *et al*. 2010
TER sites	Fachinetti *et al*. 2010
TER sites related to high transcription	Fachinetti *et al*. 2010
Sites of Rrm3 accumulation	Azvolinsky *et al.* 2009
Palindromic sequences	Lisnic *et al.* 2005
Sites of G4 quadruplex formation (predicted by sequence context *in silico*)	Capra *et al.* 2010
Sites of differential transcription in response to NMM[2]	Hershman *et al.* 2008
Regions of transcription-transcription conflicts resolved by Elc1p[3]	Hobson *et al*. 2012
Tracts of poly A or poly T $\geq$ 25 bp	SGD
Sites of RNA/DNA hybrid accumulation in *rnh1 rnh201*	Chan *et al.* 2014

[1]Most of the genomic elements were identified in Saccharomyces Genome Database (SGD) using the YeastMine tool (described: http://www.yeastgenome.org/help/video-tutorials/yeastmine). The criteria used to determine the number each elements in the genome are described in File S1. We chose to identify tracts of poly A or poly T > 25 bp based on a personal communication from Doug Koshland (University of California, Berkeley) who found an association between such tracts and the locations of R-loops.
[2]NMM is an abbreviation for N-methyl mesoporphyrin IX, a drug that binds G4 quadruplex structures.
[3]Elc1p is a protein required to remove stalled RNA polymerase II complexes (Hobson *et al*., 2012).

Table S10. Number of genomic elements represented on the microarrays

Element	# elements per genome	# elements on whole-genome array[1]	# elements in *rnh201Δ*[2]	# elements in *rnh201Δ pol2-M644L*[2]	# elements in *rnh1Δ rnh201Δ*[2]	# elements on chromosome IV array[3]
Ty elements	50	48	48	48	35	8
Solo LTRs	291	280	280	276	246	19
Centromeres	16	16	16	16	16	1
Intron-containing genes	345	331	331	323	287	24
ARS elements	352	317	317	308	275	28
tRNA genes	275	274	274	270	236	23
Long genes	306	306	306	295	259	28
Regions of high transcription	330	329	329	318	270	36
Regions of low transcription	328	312	312	303	272	20
ORFs with high GC content	115	115	115	114	96	8
High G content on the non-transcribed strand	41	41	41	40	34	1
Sites of Rbp3 accumulation in S phase	93	93	93	93	70	11
TER sites	71	71	71	69	69	3
TER sites related to high transcription	58	58	58	56	56	3
Sites of Rrm3 accumulation	115	112	112	112	103	6
Palindromic sequences	611	589	589	573	517	51
Sites of G4 quadruplex formation (predicted by sequence context *in silico*)	636	543	543	536	480	20
Sites of differential transcription in response to NMM	114	107	107	105	96	7
Regions of transcription-transcription conflicts resolved by Elc1.	144	144	144	137	125	13

Tracts of poly A or poly T ≥ 25 bp	43	41	41	40	36	3
Sites of RNA/DNA hybrid accumulation in *rnh1 rnh201*	163	129	129	127	112	10

[1]In our analysis, we examined twenty-one types of genomic elements. The references for the locations of these elements are described in Table S9 and in File S1. The total size of the yeast genome, as presented in SGD, is about 12.1 Mb. Since this calculated size counts only two of the approximately 150 ribosomal rRNA genes, there are about 12.1 Mb of single-copy yeast sequences. Our whole-genome microarrays omit most repetitive sub-telomeric repeats. 11.6 Mb of the genome are represented on our whole-genome arrays (details in Dataset S1 of Song *et al*., 2014). The coordinates and sequences of all oligonucleotides on the whole-genome arrays are in Table S5 of St. Charles *et al.* (2012).

[2]In several of the mutant strains, LOH events existed in the starting strains. These regions were omitted from the analysis. The summary of these omissions is: 1) *rnh201*Δ (no pre-existing LOH events, therefore, no omissions necessary), 2) *rnh201*Δ *pol2-M644L* (sequences distal to SGD coordinate 592645 on chromosome XIII were homozygous and deleted from the analysis; 11.3 Mb were in the remaining analysis), and 3) *rnh1 rnh201* (most of these strains had terminal LOH events beginning at SGD coordinate 1263027 on chromosome IV, and coordinate 447834 on chromosome XII; the remaining portion of the genome was about 11.0 Mb). The numbers of elements in each strain were corrected for these deletions.

[3]The chromosome IV-specific microarrays monitor SNPs from *CEN4* to near the right telomere of chromosome IV (coordinate 1511708), a region of about 1.1 Mb). The locations of SNPs on the chromosome IV-specific arrays are in Table S9 of St. Charles and Petes (2013).

Table S11. Association of LOH events in sub-cultured strains with various genomic elements.

Genome Feature	*rnh201Δ*			*rnh201Δ pol2-M644L*			*rnh1Δ rnh201Δ*		
	Exp inside: outside	Obs inside: outside	p-value	Exp inside: outside	Obs inside: outside	p-value	Exp inside: outside	Obs inside: outside	p-value
Ty elements	8:904	7:905	0.863	5:811	6:810	0.823	9:761	5:765	0.240
Solo LTRs	33:5287	24:5296	0.138	22:4670	12:4680	0.042	51:5779	66:5764	0.041
Centromeres	2:302	0:304	0.286	1:271	3:269	0.133	3:349	3:349	1.000
Intron- containing genes	39:6250	30:6259	0.173	26:5465	22:5469	0.493	60:6738	59:6739	1.000
ARS elements	37:5986	31:5992	0.362	25:5211	19:5217	0.269	57:6367	54:6370	0.729
tRNA genes	32:5174	22:5184	0.092	22:4568	16:4574	0.238	50:5626	48:5628	0.823
Long genes	46:5768	45:5769	0.920	31:4984	35:4980	0.527	72:6220	67:6225	0.597
Genes with high transcription	39:6212	35:6216	0.578	25:5381	25:5381	1.000	58:6564	63:6559	0.554
Genes with low transcription	37:5891	37:5891	1.000	24:5127	27:5124	0.610	57:6367	51:6373	0.467
ORFs with high GC content	14:2171	9:2176	0.227	9:1929	15:1923	0.066	20:2246	14:2252	0.218
High G content on the non-transcribed strand	5:774	4:775	0.823	3:677	1:679	0.387	7:763	9:761	0.572
Sites of Rbp3 accumulation in S phase	11:1756	11:1756	1.000	7:1574	6:1575	0.842	16:1766	20:1762	0.377
TER sites	8:1341	5:1344	0.377	6:1167	8:1165	0.538	18:1500	15:1503	0.554
TER sites related to high transcription	9:1093	5:1097	0.242	6:946	5:947	0.842	14:1218	11:1221	0.498
Sites of Rrm3 accumulation in S phase	13:2115	17:2111	0.330	9:1895	12:1892	0.406	21:2311	12:2320	0.063
Palindromic sequences	70:11121	80:11111	0.254	46:9695	42:9699	0.603	110:12320	85:12345	0.008
Sites of G4 quadruplex formation (predicted *in silico*)	64:10253	45:10272	0.020	43:9069	47:9065	0.597	99:11121	73:11147	0.010
Regions of differential transcription in response to NMM	13:2020	12:2021	0.888	8:1777	5:1780	0.377	20:2246	15:2251	0.313
Regions of transcription-transcription conflicts resolved by Elc1	17:2719	18:2718	0.920	11:2318	18:2311	0.050	27:2987	25:2989	0.777
Regions of RNA/DNA hybrid accumulation in the *rnh1 rnh201* mutant	15:2436	14:2437	0.888	10:2149	13:2146	0.427	24:2660	28:2656	0.475
Poly A or poly T tracts $\geq$ 25 bp	5:774	8:771	0.262	3:677	2:678	0.777	8:850	6:852	0.597

The details of the association analysis are described in File S1. In brief, we summed the amount of sequences inside the LOH association windows and the amount of sequences located outside of those windows for all isolates of the individual mutant strains. Based on the total number of elements examined by the array (Table S10), we calculated the number of elements expected within and outside of those windows; this information is summarized in the column labeled “Exp inside:outside.” We then counted the elements within and outside of the association windows (Column “Obs inside:outside”). The observed and expected values were compared using Chi-square “Goodness of Fit” test. Because of the multiple comparisons, we then applied the Benjamini-Hochberg correction to the data (Benjamini and Hochberg, 1995). Following this correction, none of the LOH events in any of the strains were significantly associated with any of the tested genomic elements.

Table S12. Association of LOH events in sectored colonies with various genomic elements on the right arm of chromosome IV

	*rnh201*Δ			*rnh1*Δ *rnh201*Δ		
Genome Feature	**expected inside: outside**	**observed inside: outside**	**p-value**	**expected inside: outside**	**observed inside: outside**	**p-value**
Ty elements	3:165	2:166	0.777	3:109	4:108	0.777
Solo LTRs	5:394	3:396	0.498	7:259	10:256	0.340
Centromeres	0:21	0:21	1.000	0:14	0:14	1.000
Intron-containing genes	7:497	8:496	0.841	8:328	12:324	0.210
ARS elements	8:580	9:579	0.862	10:382	7:385	0.420
tRNA genes	6:477	5:478	0.585	8:314	9:313	0.862
Long genes	10:578	15:573	0.150	11:381	13:379	0.647
Genes with high transcription	10:746	8:748	0.632	13:491	17:487	0.327
Genes with low transcription	6:414	4:416	0.532	7:273	4:276	0.340
ORFs with high GC content	2:166	1:167	0.718	3:109	1:111	0.380
Regions with high G content on the non-transcribed strand	0:21	0:21	1.000	0:14	1:13	1.000
Sites of Rbp3 accumulation in S phase	3:228	3:228	1.000	4:150	3:151	0.806
TER sites	1:62	1:62	1.000	1:41	1:41	1.000
TER sites related to high transcription	1:62	1:62	1.000	1:62	1:62	1:62
Sites of Rrm3 accumulation in S phase	2:124	1:125	0.718	2:82	1:83	0.718
Palindromic sequences	14:1057	17:1054	0.498	18:696	27:687	0.043
Sites of G4 quadruplex formation	6:414	6:414	1.000	7:273	9:271	0.566
Regions of differential transcription in response to NMM	2:145	1:146	0.718	2:96	2:96	1.000
Regions of t ranscription-transcription conflicts resolved by Elc1	4:269	4:269	1.000	5:177	8:174	0.256
Sites of RNA/DNA hybrid accumulation in the *rnh1 rnh201* mutant	3:207	4:206	0.777	4:136	7:133	0.206
Poly A or poly T tracts $\geq$ 25bp	1:62	1:62	1.000	1:41	2:40	0.610

The associations were examined by the same methods described in Table S11, except analysis was limited to the right arm of chromosome IV. None of the associations was statistically significant after applying the correction for testing multiple samples.

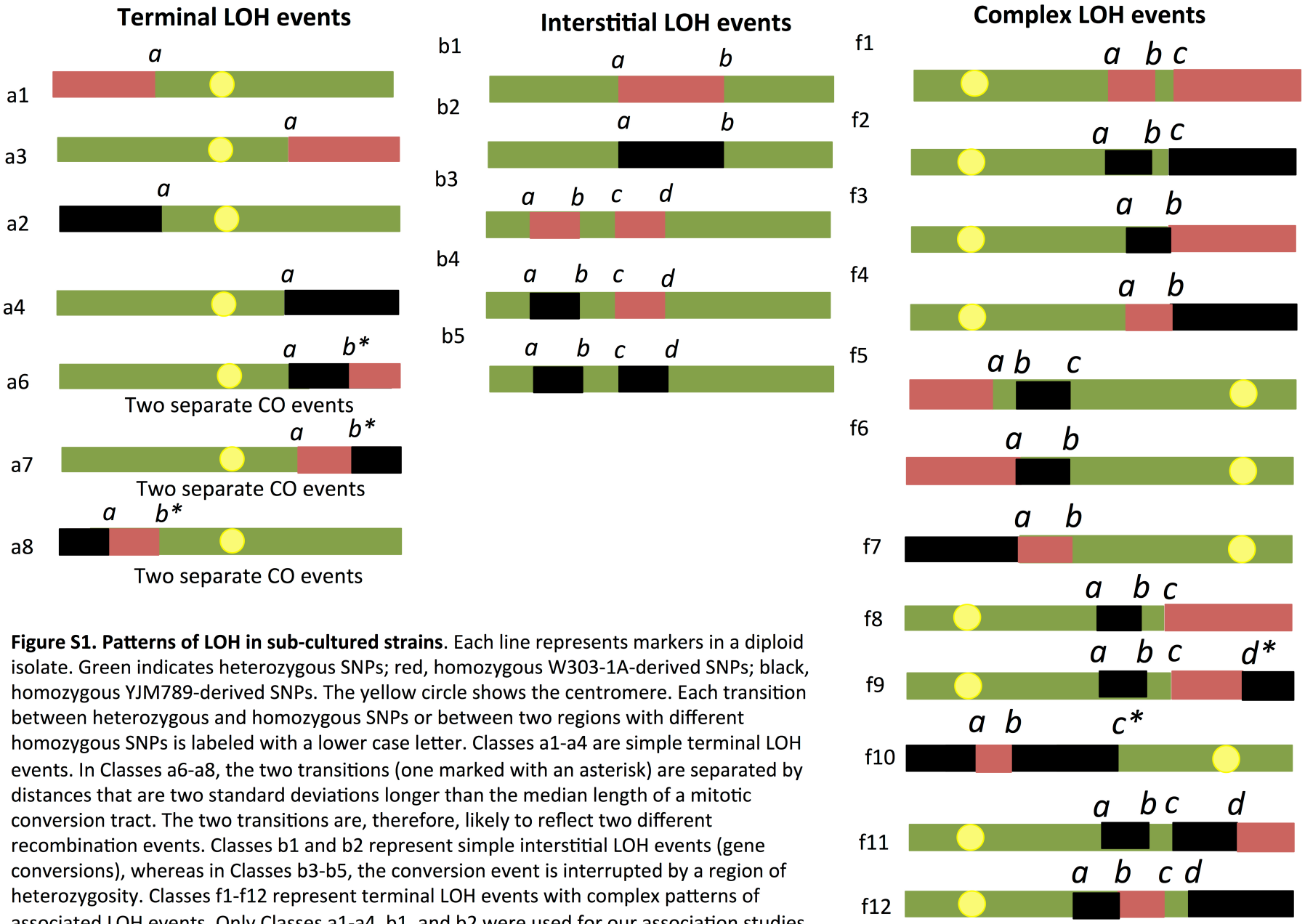

Figure S1. Patterns of LOH in sub-cultured strains. Each line represents markers in a diploid isolate. Green indicates heterozygous SNPs; red, homozygous W303-1A-derived SNPs; black, homozygous YJM789-derived SNPs. The yellow circle shows the centromere. Each transition between heterozygous and homozygous SNPs or between two regions with different homozygous SNPs is labeled with a lower case letter. Classes a1-a4 are simple terminal LOH events. In Classes a6-a8, the two transitions (one marked with an asterisk) are separated by distances that are two standard deviations longer than the median length of a mitotic conversion tract. The two transitions are, therefore, likely to reflect two different recombination events. Classes b1 and b2 represent simple interstitial LOH events (gene conversions), whereas in Classes b3-b5, the conversion event is interrupted by a region of heterozygosity. Classes f1-f12 represent terminal LOH events with complex patterns of associated LOH events. Only Classes a1-a4, b1, and b2 were used for our association studies.

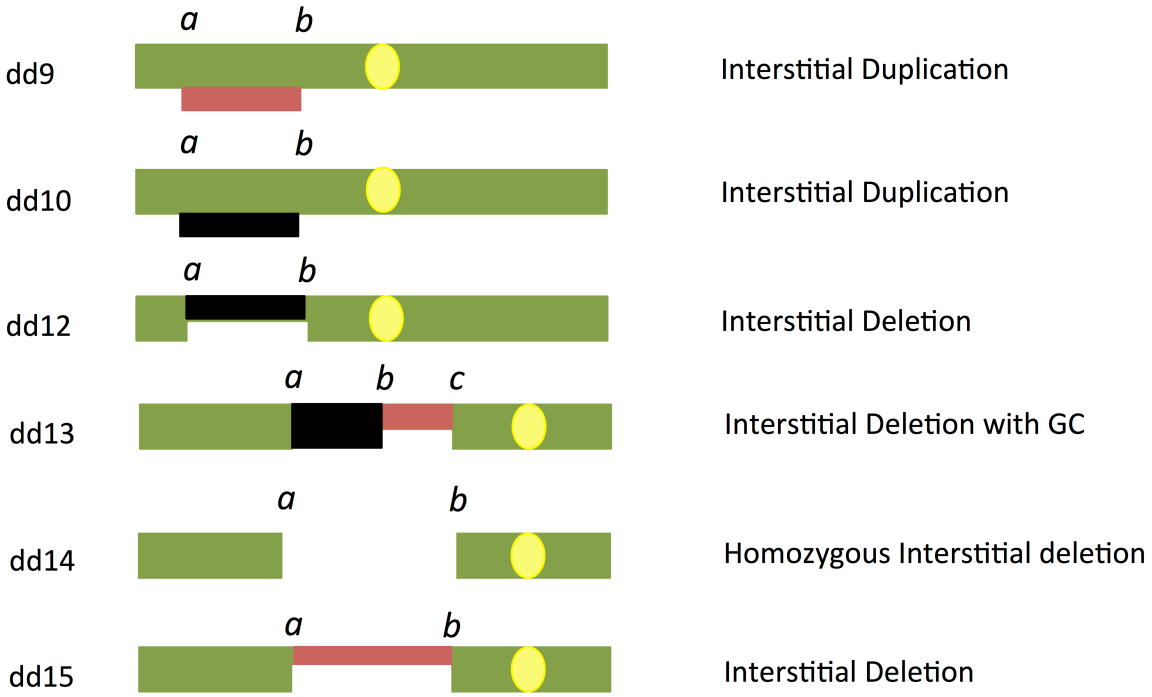

Figure S2. Deletions and duplications in sub-cultured strains. This diagram shows the patterns of deletions and duplications in diploid isolates. As in Fig. S1, the green line indicates heterozygous SNPs, and yellow circles show the centromere. The deletion or duplication is shown as line that is half as wide as the green lines. The Classes dd9 and dd10 show interstitial duplications of W303-1A-derived and YJM789-derived SNPs, respectively. The Class dd12 shows an interstitial deletion in which W303-1A-derived SNPs were removed. The coordinates for the transitions are in Table S6.

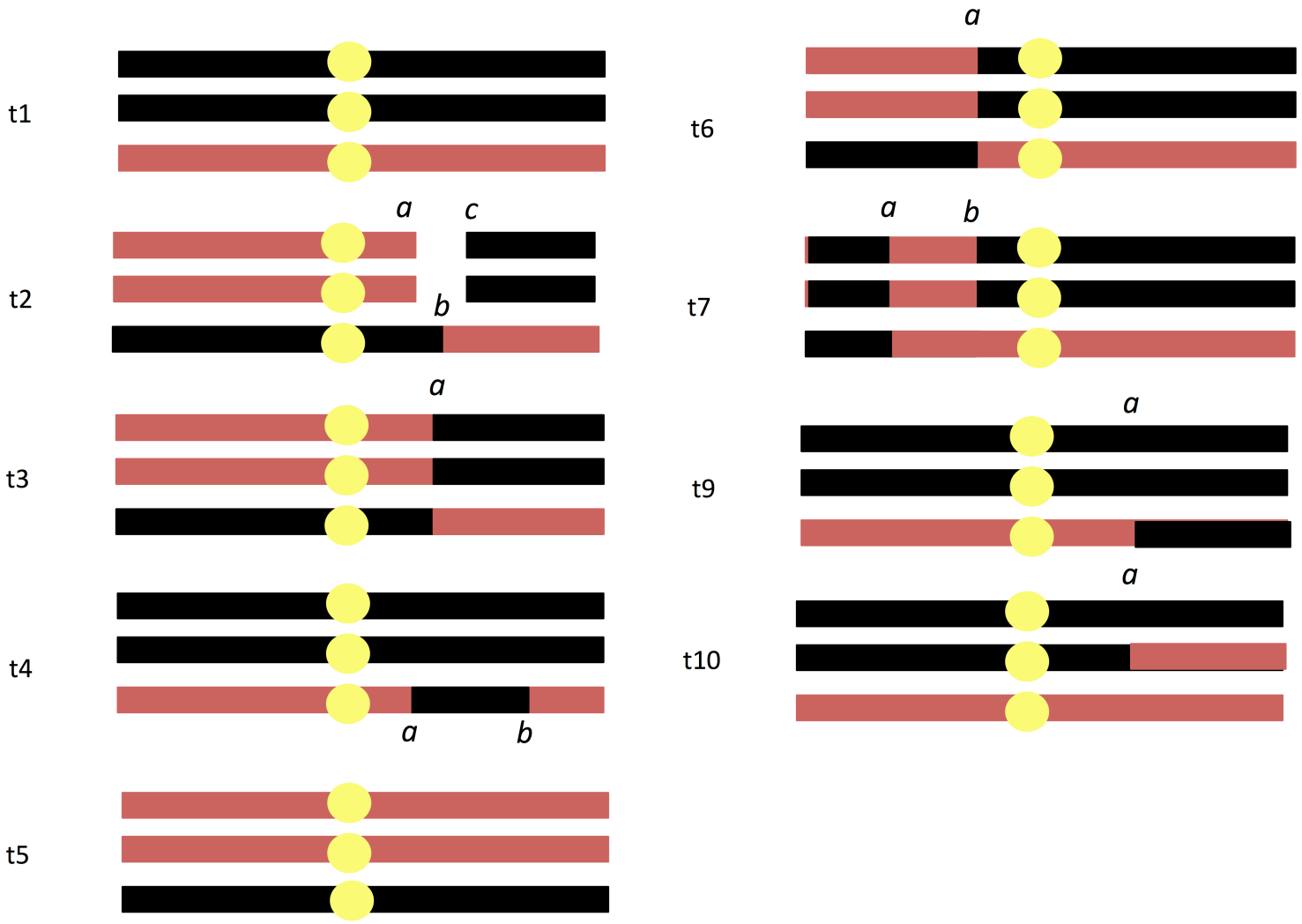

Figure S3. Aneuploidy events in sub-cultured strains. Trisomic, but not monosomic, aneuploid events were observed in our studies. For each chromosome, we indicate whether the strain has W303-1A-derived SNPs (red) or YJM789-derived SNPs (black). Note that many of these aneuploid events are associated with recombination on one or more chromosomes.

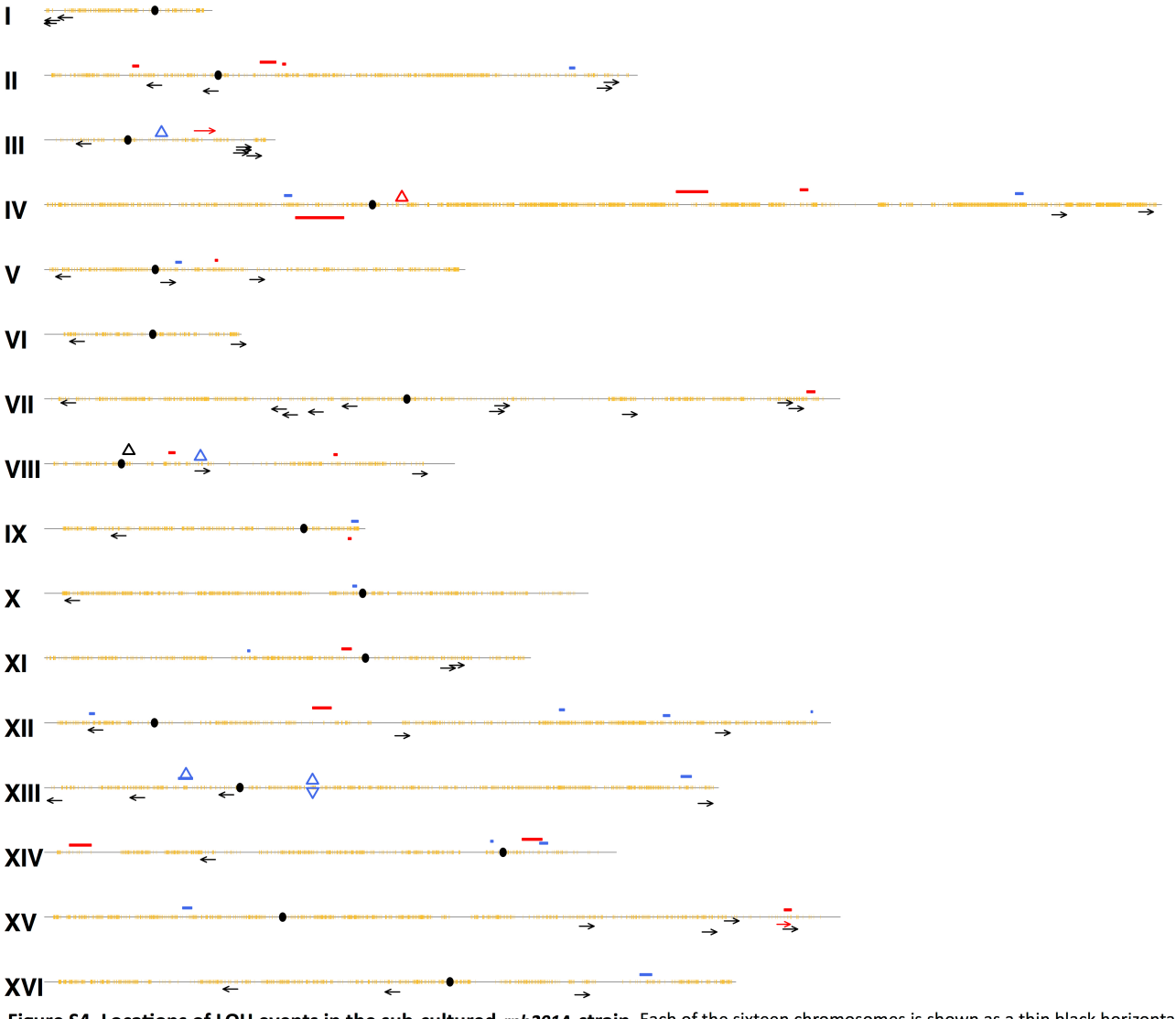

Figure S4. Locations of LOH events in the sub-cultured *rnh201Δ* strain. Each of the sixteen chromosomes is shown as a thin black horizontal line with SNPs shown as very short vertical yellow lines. The centromeres are represented by black ovals. Red and blue bars show regions of interstitial LOH in which the W303-1A-derived SNPs became homozygous and the YJM789-derived SNPs became homozygous, respectively. Black arrows indicate the positions of terminal LOH events that were unassociated with a gene conversion event. Red arrows and blue arrows indicate terminal LOH events that were associated with a conversion that made W303-1A-derived SNPs and YJM789-derived SNPs homozygous, respectively. Triangles indicate deletions (red for a deletion of W303-1A-derived sequences and blue for a deletion of YJM789-derived sequences) and inverted triangles indicate duplications (same color code as for deletions).

Figure S5. Location of LOH events in the sub-cultured *rnh1Δ rnh201Δ* strain. The mapped events are shown with the same code as Fig. S4

Figure S6. Location of LOH events in the sub-cultured *rnh201Δ pol2-M644L* strain. The mapped events are shown with the same code as in Fig. S4

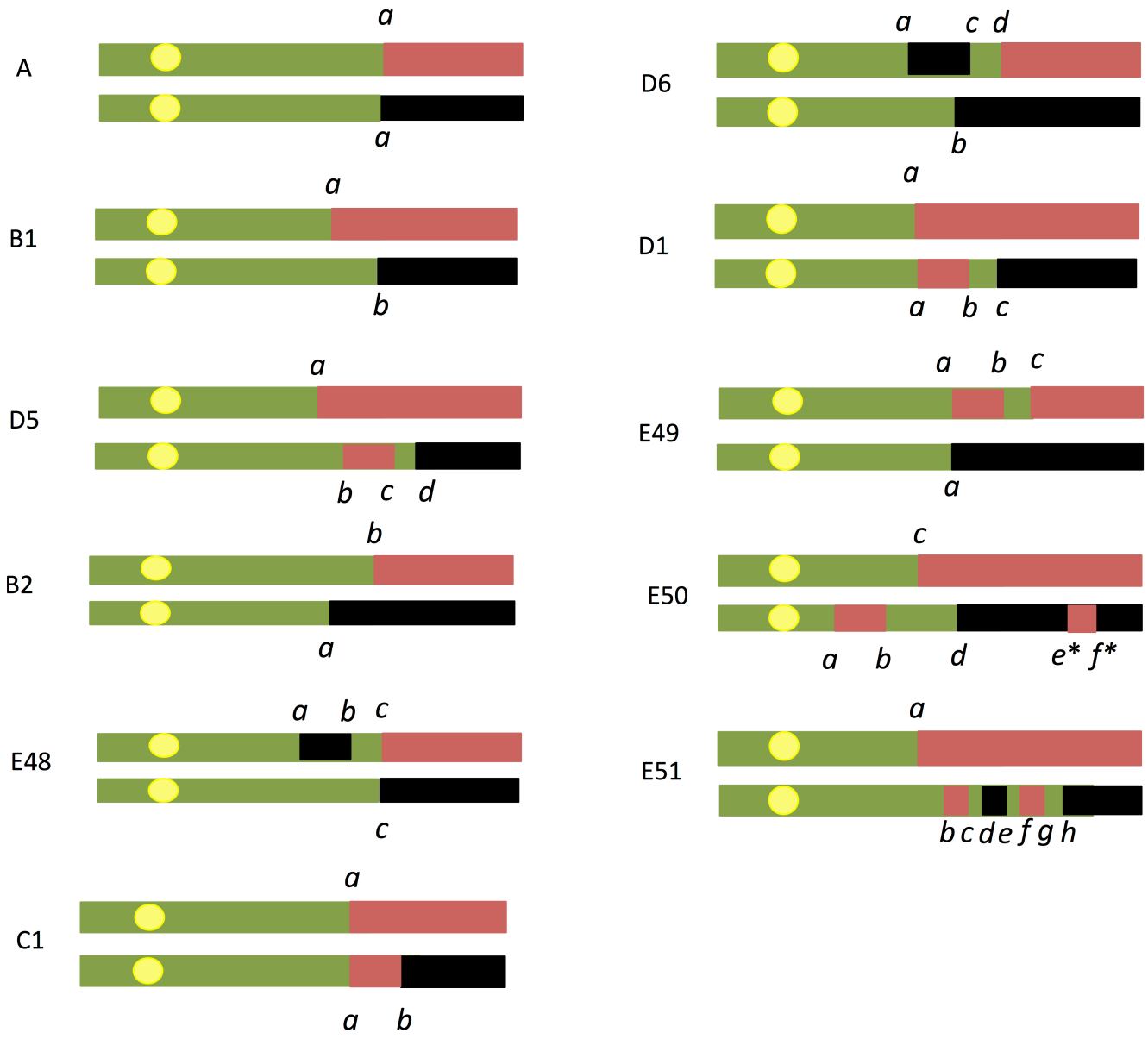

Figure S7. Patterns of LOH in sectored colonies. In this depiction, each sectored colony is represented by a pair of lines with the red sector shown as the top line. We use the same color code for heterozygous and homozygous regions as in Fig. S1. Other features of these patterns are described in the main text.